AD_____

Award Number: DAMD17-01-1-0160

TITLE: Metastic Progression of Breast Cancer by Allelic Loss on Chromosome 18q21

PRINCIPAL INVESTIGATOR: Sam Thiagalingam, Ph.D.

CONTRACTING ORGANIZATION: Boston University
Boston, Massachusetts 02118

REPORT DATE: September 2003

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 2003	Annual Summary (1 Sep 2002 - 31 Aug 2003)

4. TITLE AND SUBTITLE
Metastic Progression of Breast Cancer by Allelic Loss on Chromosome 18q21

5. FUNDING NUMBERS
DAMD17-01-1-0160

6. AUTHOR(S)
Sam Thiagalingam, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Boston University
Boston, Massachusetts 02118

E-Mail: samthia@bu.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Mutations in the Smad2 and Smad4 genes, localized to chromosome 18q, a region frequently deleted in advanced cancers is rare in breast cancer unlike pancreatic, colon, lung, and ovarian cancers. However, the fact that 18q loss has been predominantly associated with the advanced stage of cancers suggests that the genes inactivated by this specific alteration or other genes in the same pathway targeted for the inactivation could be associated with the conversion of benign tumors to malignancy and metastatic progression of breast cancer. Our analysis showed that 30% of the breast cancers exhibit loss of Smad8 expression, making it on of the highly valued markers similar to Her/neu. We also provide the first direct evidence that the silencing of gene expression occurs via DNA hypermethylation of the Smad8 gene and it could be an important event in breast cancer progression and metastatic. Therefore, Smad8 has the potential to become a key target for the development of diagnostic, prognositic, and therapeutic strategies to combat breast cancer. We are continuing to characterize and identify the mediator and effector genes, which regulate metastatic progression of breast cancer upon inactivation of the Smad signaling pathways.

14. SUBJECT TERMS
Smad, signaling, metastasis, breast cancer, TEGD, Chromosome 18, TGFβ

15. NUMBER OF PAGES
66

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

ANNUAL REPORT OF THE USAMRMC FUNDED ACTIVITY

Title of the grant: Metastatic progression of breast cancer by allelic loss on chromosome 18q21.

1. Introduction/ Project Overview/ Scientific Progress and future directions:

The majority of molecular genetic studies on breast cancer have focused on familial predisposition and there has been a lack of serious effort to understand the molecular basis of the involvement of genetic determinants in the progression to metastatic cancer. The fact that 18q loss has been predominantly associated with the advanced carcinoma stage of cancers suggests that the genes inactivated by this specific alteration or other genes in the pathway targeted for the inactivation could be associated with the conversion of benign tumors to malignancy and metastatic progression of breast cancer. However, unlike in pancreatic, colon, lung and ovarian cancers, the lack of mutations in breast cancer in the *Smad2* and *Smad4* genes localized to chromosome 18q, strongly supports the existence of alternate target genes in breast cancer.

Disabling Smad signaling in cancer has become increasingly recognized as an important step that affects processes such as loss of growth inhibition, promotion of angiogenesis and metastasis and the epithelial mesenchymal transition (1, 2). Our survey of the various *Smad* genes has provided the first clues in identifying the *Smad8* gene as an important target for loss of expression in nearly 30% of breast cancers which we believe is a significant finding as even the most celebrated tumor marker, *HER/neu* gene amplification, also occurs in about 20%-30% breast cancer cases (3). We report here that we have extended these initial observations to demonstrate that the inactivation of the *Smad8* gene leading to loss of its expression is mediated by epigenetic DNA methylation (4). It still remains to be determined whether *Smad8* inactivation could also be an alternate target for *Smad2* or *Smad4* inactivation.

Furthermore, our investigation of the potential role of Smad4 inactivation revealed that the gene expression pattern in cell culture models that lack Smad4 could favor angiogenesis/ metastasis, which is further enhanced by TGFβ and hypoxia. We are continuing to characterize these cell culture models to identify the mediator and effecter genes, which regulate metastatic progression of breast cancer upon inactivation of the Smad4 signaling pathway.

2. Modified tasks that were approved following the first annual report, their expansion to set specific goals, summary of findings and future directions:

We have further expanded the tasks 1 & 2 to set specific goals incorporating our recent findings that are aimed at increasing the understanding of the implications of the role of chromosome 18q loss in the molecular basis of metastatic breast cancer.

Task 1. Determination and identification of genetic and epigenetic alterations in known and novel *Smad*s as potential target genes and the elucidation of their implications to metastatic breast cancer.

We have employed a novel technique known as TEGD (targeted expressed gene display) to identify that the loss of *Smad8* gene expression is the major *Smad* gene target for inactivation in breast cancer. We also demonstrate that the epigenetic silencing of *Smad8* expression by DNA hypermethylation directly correlates with loss of *Smad8* expression. We are in the process of molecular cloning wildtype and defective *Smad8* genes to carry out adding back experiments to further understand the role of the Smad8 inactivation in cancer. We also plan to determine whether Smad8 inactivation could be an alternate target for the inactivation of *Smad2* or *Smad4* or deregulation of Smad7.

Task 2. Identification and elucidation of the roles of alternate target genes involved in the Smad4 signaling pathway.

We have made progress in establishing isogenic cell culture model systems that are proficient and deficient in Smad4 expression. Our preliminary data strongly support that the *Smad4* defect could be a critical contributor in the gene expression pattern that favor angiogenesis/ metastasis. Interestingly, these events are highly favored by TGFβ and hypoxia consistent with the conditions that promote advanced cancer. Further characterization and identification of mediator and effecter genes that promote angiogenesis/ metastasis under these conditions and the identification of potential cofactors that could interact with Smad4 and hence could be alternate target(s) for inactivation in breast cancer are in progress.

Task 3. Evaluation of candidate target genes.

This task remains unmodified and would begin once we have identified legitimate target genes in Tasks 1 & 2.

We have made substantial progress towards not only the identification of the major *Smad* gene target (*Smad8*) but also the molecular basis of its loss of function in breast cancer. We have also established the model systems and conditions that should aid us in the discovery of alternate targets in Smad4 signaling as well as the effecter/ mediator genes that are involved in the genesis/ progression of metastatic breast cancer. We believe that these studies could provide important insights into the molecular basis of breast cancer metastasis leading to better diagnosis, prognosis and therapy of the disease.

3. Body: Procedures and progress report:

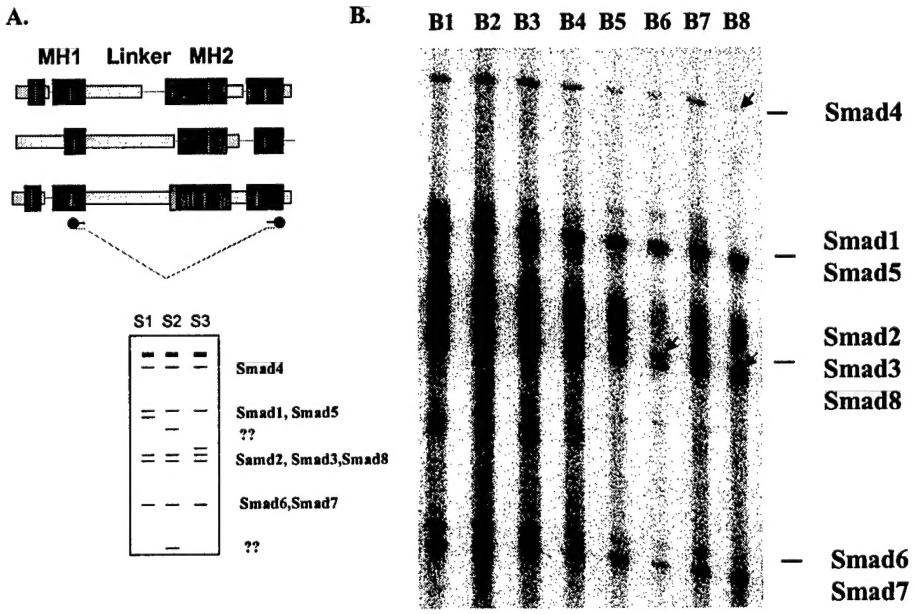

Figure 1. Targeted expressed gene display (TEGD).

A. Schematic representation of TEGD for the Smad family of genes.
MH1 and MH2 indicate highly homologous regions in the amino acid as well as DNA sequence among the various *Smad* gene family members. The forward and reverse primers for PCR amplification of the cDNA were designed in the conserved regions as indicated. The radiolabeled PCR products were analyzed by denaturing acrylamide gel electrophoresis. B. PCR products for *SMADs* using degenerate primers were analyzed by TEGD. Lanes B1-8 correspond to PCR products generated using cDNA templates from the normal mammary gland cells (B1) and tumor or cell line (B2-8) samples. B8 is a cell line (MDAMB468) with a homozygous deletion for Smad4 and serves as an internal control. The arrows point to distinct PCR products that were abnormal compared to the normal control. The positions of various Smad genes and their variants as identified from sequence analysis are indicated on the right panel.

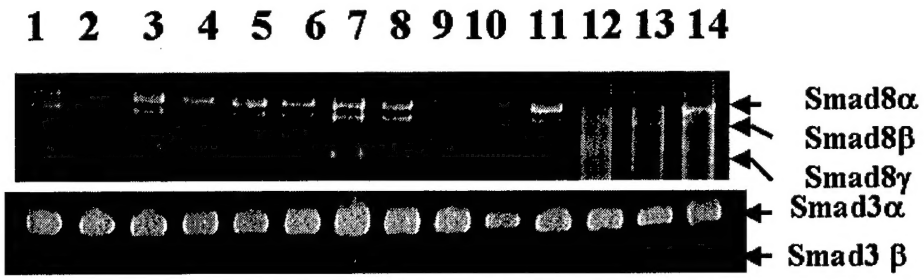

Figure 2. Semi-quantitative RT-PCR analysis of Smad8 expression in breast cancer.
Total RNA was prepared using the Trizol method from the indicated breast cancer specimens and analyzed by RT-PCR. Lane 1 is a normal breast sample and lanes 2-4 and 12-14 are primary tumor samples, 5-11 are cell lines. Smad8α, Smad8β and Smad8γ are three of the major differentially spliced forms of Smad8 which correspond to the full-length, deletion of exon 2, and deletions of exons 2&3, respectively. Analysis of the *Smad3* gene is used for normalization and quantitation of *Smad8*.

The *Smad* family of genes has highly homologous amino acid sequences at their N- and C-terminal regions (MH1 and MH2 respectively), which are separated by a highly divergent linker region rich in proline, serine and threonine (1; Figure 1A). We have effectively exploited TEGD as a tool to identify *Smad8* gene as a critical target for loss of function due to down regulation of gene expression in breast cancer (Figures 1A & 1B). Subsequent analysis of the *Smad8* gene using gene specific primers by semi quantitative RT-PCR in breast and other cancers showed loss of expression in nearly 31% (11/35) of breast cancers (Figure 2). We believe that it is a significant finding as even the most celebrated tumor marker for breast cancer, the *HER/neu* gene amplification, occurs in about 20%-30% breast cancer cases (3).

We have decided to extend these observations to investigate potential mechanisms for the loss of *Smad8* gene expression due to the high level of significance of this alteration in breast cancer and its potential implication to the design of diagnostic and therapeutic strategies. Since our analysis of chromosomal deletions was negative, we considered epigenetic silencing of gene expression due to DNA methylation and associated chromatin modification (4). DNA sequence analysis of the bisulfite treated genomic DNA revealed that CpG islands localized to nucleotides 3541028 to 35410583 (Chromosome 13q12-14 (on the reverse strand between *Rb* and *BRCA2*; UCSC genome browser http://genome.ucsc.edu) in the first intron of the *Smad8* gene is only methylated in cancers that exhibited loss of expression (data not shown). We confirmed these observations using methylation specific PCR (MSP) using primers designed to these corresponding differentially methylated regions and the results were consistent with the earlier observations that the *Smad8* gene is silenced in breast cancer due to DNA hypermethylation affecting CpG islands in the first intron of the *Smad8* gene (Figure 3A).

Furthermore, the physiological significance of the role of DNA hypermethylation in Smad8 gene silencing was established with the ability to recover gene expression upon treatment with 5'-aza-2'-deoxycytidine (5Aza-dC; a DNA demethylating agent) in cell lines that were previously determined as exhibiting DNA hypermethylation mediated gene silencing of *Smad8* (Figure 3B). These observations strongly support the loss of Smad8 expression in breast cancer is primarily mediated by hypermethylation of cis-regulatory CpG islands of the gene.

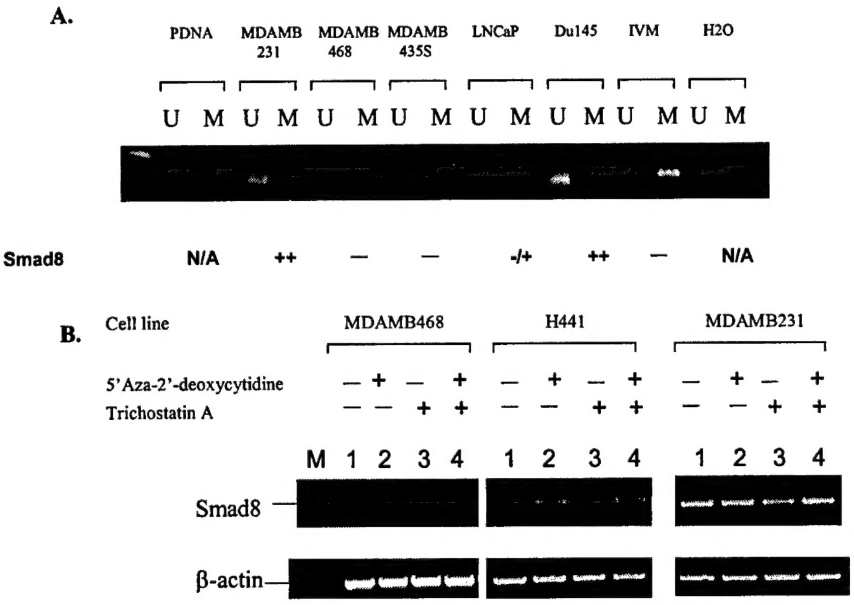

Figure 3. Epigenetic gene silencing of the *SMAD8* gene by altered DNA methylation patterns.
A. MSP (Methylation specific PCR) analysis of the CpG islands of intron 1 of the *Smad8* gene in the indicated breast (MDAMB231, MDAMB468, MDAMB435S) and prostate (LNCaP, Du145) cancer cell lines that are either proficient (+) or deficient(-) in *Smad8* expression. Placental DNA (PDNA) and *in vitro* methylated DNA (IVM) serve as negative and positive controls. Lanes U and lanes M indicate the presence of unmethylated and methylated templates, respectively.
B. The indicated cell lines were treated with 1-5μM 5-AZA-dC for 7 days or with 300μM TSA for 24hrs. To assess the effect of both 5-AZA-dC and TSA simultaneously, cells were exposed sequentially for 7 days to 5-AZA-dC and subsequently to 300 μM TSA for an additional 24 hrs. Total RNA and genomic DNA were isolated and *Smad8* expression and DNA hypermethylation were determined by RT-PCR and MSP analysis (data not shown), respectively.

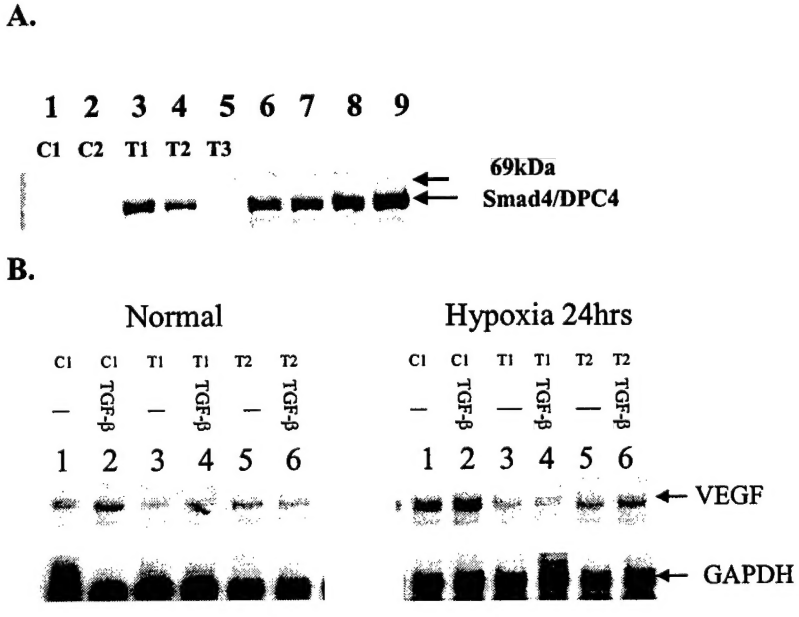

Figure 4. Relationship between *Smad4* status and the expression of VEGF.
A. *Western blotting was used to screen for stable cell lines that constitutively express Smad4 and corresponding isogenic controls that have*

integrated the empty vector. Lanes 1-5 correspond to derivatives of a colon cancer cell line (CC1) with Smad4-/- stably transfected with empty vector (1&2) or pCMV-Smad4 (3-5). Lanes 6-9 are a breast cancer cell line, BR05(Smad4^{mt}) stably transfected with an empty vector (6&8) or TGFβ-RII receptor (7) or pCMV-Smad4 (9). Please note that the clone in lane 5 is a false positive. B. *Effect of over-expression of the Smad4 gene on VEGF.* Total RNA was analyzed with the RiboQuant probes (BD-PharMingen, San Diego, CA) to detect the indicated mRNAs. GAPDH was included as an internal control. C1, T1 & T2 are stable transfectants of CC1 Smad4-/- with empty vector (C1) or Smad4 expression vector (T1 & T2). The evaluations were made in the presence/ absence of TGFβ and under normoxic/ hypoxic conditions as indicated.

In summary, we conclude that our preliminary data provides the first direct evidence that silencing of gene expression *via* DNA hypermethylation of the *Smad8* gene could be an important event in breast cancer progression and metastasis.

Furthermore, preliminary results from the experiments to investigate the role of Smad4 in cancer metastasis are encouraging as the introduction of wild-type Smad4 into a colon cancer cell line with homozygous deletion of Smad4 exhibited a decrease in VEGF expression (Figure 4). Interestingly, the presence of TGFβ and hypoxic conditions that mimic advanced tumors elicited a significant increase in the expression of VEGF, a marker for angiogenesis/ metastasis. These studies are currently being repeated in breast cancer cell culture models.

We are planning to extend these studies to not only confirm this phenomenon with other candidate genes but also identify a wide spectrum of other critical genes important for the metastatic progression of breast cancer using the microarray (Affymetrix) technology.

Once legitimate metastasis mediator and effecter gene(s) are identified, evaluation of the status of the candidate gene(s) for inactivation/ activation in metastatic breast cancer will commence as described in the original proposal (5; Task 3).

4. Key research accomplishments:

Our study provide the first direct evidence that 30% of the breast cancers exhibit loss of Smad8 expression and makes it as one of the highly valued markers similar to *Her/neu.* Our studies also provide the first direct evidence that the silencing of gene expression *via* DNA hypermethylation of the *Smad8* gene could be an important event in breast cancer progression and metastasis. Therefore, Smad8 has the potential to become a key target for the development of diagnostic, prognostic and therapeutic strategies to combat breast cancer.

We have also identified/ generated appropriate tumor cell lines as well as experimentally developed derivative test and control cell lines as model systems to identify and isolate the metastatic breast cancer mediator and effecter genes involved in the Smad4 signaling pathway.

5. Conclusions:

(1) The loss of *Smad8* expression in breast cancers is primarily mediated by gene silencing due to epigenetic DNA methylation of regulatory regions.

(2) A combination of *Smad4* inactivation, high levels of TGFβ and hypoxic conditions could favor angiogenesis/ metastasis.

(3) The identification of target gene(s) that disable Smad4 or Smad8 signaling to promote breast cancer could potentially provide not only novel and valuable diagnostic and prognostic tumor markers but also key arsenals to combat breast cancer.

6. References:

1. Thiagalingam, S., K-h.Cheng, R. L. Foy, H. J. Lee, D. Chinnappan, and J. F. Ponte. 2002. TGFβ and its *Smad* connection to cancer. *Current Genomics* 3: 449-476.
2. Schwarte-Waldhoff I, Volpert OV, Bouck NP, Sipos B, Hahn SA, Klein-Scory S, Luttges J, Kloppel G, Graeven U, Eilert-Micus C, Hintelmann A, and Schmiegel W. 2000.

Smad4/DPC4-mediated tumor suppression through suppression of angiogenesis. *Pro. Natl. Acad. Sci.* 97: 9624-9629.

3. Zujewski, J.A. 2002. "Build Quality in"-HER2 Testing in the Real World. *J Natl Cancer Inst.* **94**: 788-789.
4. Baylin, S. B, and J. G. Herman. 2000. DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet.* **16**: 168-174.
5. Thiagalingam, S. 2001. Molecular detection of Smad4/Smad2 alterations in colorectal tumors: Colorectal Cancer Methods and Protocols. In *Methods in Molecular Medicine.* S. M. Powell (Ed) Humana Press Inc., New Jersey, 50: 149-165.

7. Scientific presentation/ publications/ patent relevant to this grant:

Seminar by Dr. Sam Thiagalingam:
Smad connection to colon cancer. Pathology Workshop, Boston University School of Medicine, August 8, 2003.

Publications:

1. Cheng, K-h., J. F. Ponte, and S. Thiagalingam. 2003. Elucidation of *Smad8* inactivation in cancer using Targeted Expressed Gene Display. *Manuscript submitted.*

2. Thiagalingam, S., K-h.Cheng, H. J. Lee, N. Mineva, and J. F. Ponte. 2003. Histone deacetylases: Unique players in shaping the epigenetic histone code, *Annal. New York Acad. Sci.* 983: 86-100.

Patent:

1. Method of determining gene expression-Targeted Expressed Gene Display – Filed provisional application (BU03-16; March 25, 2003); Boston University.

APPENDIX II -Resume

Sam Thiagalingam, MS., PhD.

Sam Thiagalingam, Ph. D.

Assistant Professor
Boston University School of Medicine
715 Albany Street, L320
Boston, MA 02118

Phone #: 617 638 6013
Fax #: 617 638 4275
E-mail: samthia@bu.edu

EDUCATION

University of Jaffna, Thirunelvely, Sri Lanka.	B.Sc. 1982	Biology
Bowling Green State University, Bowling Green, OH	M.S. 1986	Biology-Microbiology
The Johns Hopkins University, Baltimore, MD	Ph.D. 1992	Biochemistry
The Johns Hopkins Oncology Center, Baltimore, MD	1991 - 1995	PDF-Molecular Genetics/ Oncology

Graduate and Post-doctoral Advisors:

Doctoral Thesis Advisor: Dr. Lawrence Grossman, University Distinguished Service Professor of Biochemistry and Molecular Biology, The Johns Hopkins University, Baltimore, Maryland.

Post-doctoral Advisor: Dr. Bert Vogelstein, Clayton Professor of Oncology & Howard Hughes Investigator, The Sidney Kimmel Comprehensive Cancer Center, The Johns Hopkins School of Medicine, Baltimore, Maryland.

PROFESSIONAL EXPERIENCE

1981 - 1982	Assistant Lecturer, Faculty of Science, University of Jaffna, Sri Lanka
1982 - 1984	Assistant Lecturer, Faculty of Science, Eastern University, Sri Lanka
1995 - 1998	Research Associate, The Johns Hopkins University School of Medicine
1998 -	Assistant Professor of Medicine, Boston University School of Medicine
1999 -	Assistant Professor of Pathology and Laboratory Medicine, Boston University School of Medicine
2003 -	Assistant Professor of Genetics and Genomics, Boston University School of Medicine

AWARDS AND HONORS

1983	Overseas Research Students Award, CVCP of the Universities of the U. K.
1988-1991	Post Certified Student Scholarship, The Johns Hopkins University
1992-1995	Amgen Post-doctoral Fellowship
1995	Oncology Fellow Research day Poster Award
1995-	Sterling Who's Who
1997-	Who's Who in the East
1999	American Lung Association Research Award
1999	American Cancer Society Institutional Research Grant Award
2000-2002	The Dolphin Trust Investigator (New Investigator Award), The Medical Foundation, MA
2001	Burroughs Wellcome Fund New Investigator in Toxicological Sciences Award (Finalist)
2001-2005	Career Development Award, Department of Defense BCRP, U.S. Army MRMC

PROFESSIONAL ORGANIZATIONS

American Association for Cancer Research	-	Member
American Society of Human Genetics	-	Member
American Society for Biochemistry and Molecular Biology	-	Member
American Association for the Advancement of Science	-	Member
Tamil Academy for the Advancement of Science	-	Vice President

EDITORIAL RESPONSIBILITIES

Editorial boards: Cancer Biology and Therapy, Journal of Molecular Biology and Biotechnology.

Ad hoc reviewer: Cancer Research, Cell Growth & Differentiation, Cancer Detection and Prevention, Cancer Epidemiology, Biomarkers and Prevention, Clinical Genetics, The American Journal of Physiology - Cell Physiology, Physiological Genomics, Proceedings of the National Academy of Sciences and Neuropsychiatric Genetics-The American Journal of Medical Genetics.

GRANT REVIEW

Ad hoc reviewer: R03 Awards, NIDA, NIH

Ad hoc reviewer: Alzheimer's Association

Peer Review Panel Member (Cell Biology): 2003 Breast Cancer Research Program/ USAMRMC

PATENTS

1. *Mad*-related genes in the human – US Patent No: 6,255,464 (07/03/2001); The Johns Hopkins University.

2. Method of determining gene expression-Targeted Expressed Gene Display – Filed provisional application (BU03-16; March 25, 2003); Boston University.

BIBILIOGRAPHY

Number of citations (Institute for Scientific Information) as of 09/28/03: 2,897

Cheng, K-h., J. F. Ponte and **S. Thiagalingam.** 2003. Elucidation of *Smad8* inactivation in cancer using Targeted Expressed Gene Display. *Manuscript submitted.*

Ponte, J. F., M. Levy, J. S. Brody and **S. Thiagalingam**. 2003. Modulation of p53 by hBUB1 is involved in the mitotic checkpoint. *Manuscript submitted.*

Kugathasan, S, Cheng K-h, Johnson N, Ogawa H, Binion DG and **S. Thiagalingam.** 2003. Inhibition of mucosal SMAD signaling is associated with the progression from early to late Crohn's disease (CD). *Manuscript in preparation.*

Ponte, J. F., K-h.Cheng, J. S. Brody and **S. Thiagalingam.** 2003. Detection of genetic damage in pre-neoplastic cells: A non-invasive diagnostic procedure. *Manuscript in preparation.*

Kugathasan S, Cheng K-h, Johnson NA, Ogawa H, Rafiee P, Binion DG, and **S. Thiagalingam.**

2003. Inactivation of mucosal SMAD signaling is associated with the progression from early to late Crohn's disease (Cd). Gastroenterology , **124 (4)**: Suppl. A313-A313

Thiagalingam, S., K-h.Cheng, H. J. Lee, N. Mineva, and J. F. Ponte. 2003. Histone deacetylases: Unique players in shaping the epigenetic histone code, *Annal. New York Acad. Sci.* **983**: 86-100.

Thiagalingam, S., K-h.Cheng, R. L. Foy, H. J. Lee, D. Chinnappan, and J. F. Ponte. 2002. TGF-β and its *Smad* connection to cancer. *Current Genomics* **3**: 449-476.

Thiagalingam, S., R. L. Foy, K-h.Cheng, H. J. Lee, A. Thiagalingam, and J. F. Ponte. 2002. Loss of heterozygosity as a predictor to map tumor suppressor genes in cancer: molecular basis of its occurrence. *Current Opinion in Oncology* **14(1)**: 65-72.

Thiagalingam, S., S. Laken, J. K. V. Willson, S. Markowitz, K. W. Kinzler, B. Vogelstein and C. Lengauer. 2001. The mechanisms underlying losses of heterozygosity in human colorectal cancers. *Proc. Natl. Acad. Sci.* USA. 98: 2698-2702.

Thiagalingam, S. 2001. Molecular detection of Smad4/Smad2 alterations in colorectal tumors: Colorectal Cancer Methods and Protocols. In *Methods in Molecular Medicine*. S. M. Powell (Ed) Humana Press Inc., New Jersey, **50**: 149-165.

Fink, S. P., S. E. Swindler, R. S. Brady, J. D. Lutterbaugh, J. Massague, **S. Thiagalingam,** K. W. Kinzler, B. Vogelstein. J. K. V. Willson, and S. Markowitz. 2000. TGF-β induced inhibition in a Smad4 mutant colon adenoma cell line. *Cancer Res.* **61**: 256-260.

Thiagalingam, S., Sau, K., Cheng, K-h., Foy, R. L., and J.F. Ponte. 2000. Identification of upstream kinase regulators of p53 function. *FASEB Journal* **14(8):** 1331

Grady, W. M., L. L. Myeroff, S. E. Swindler, A. Rajput, **S. Thiagalingam**, J. D. Lutterbaugh, A. Neumann, M. G. Brattain, J. Chang, S.-J. Kim, K. W. Kinzler, B. Vogelstein. J. K. V. Willson, and S. Markowitz. 1999. Mutational inactivation of TGF-β receptor type II in microsatellite stable colon cancers. *Cancer Res.* 59: 320-324.

Hermeking, H, C. Lengauer, K. Polyak, T.-C. He, L. Zhang, **S. Thiagalingam**, K. W. Kinzler and B. Vogelstein. 1997. 14-3-3-σ mediates a p53 induced G2 arrest. *Molecular Cell* 1: 3-11.

He, T.-C, L. T.Costa and **S. Thiagalingam.** 1997. Homeosis and Polyposis: A Tale from the Mouse. *BioEssays* **19**: 551- 555.

Riggins, J. G., K. W. Kinzler, B. Vogelstein and **S. Thiagalingam.** 1997. Frequency of *Smad* gene mutations in human cancers. *Cancer Res.* 57: 2578-2580.

Riggins, J. G., **S. Thiagalingam**, E. Rozenblum, C. L. Weinstein, S. E. Kern, S. R. Hamilton, J. K. V. Willson, S. Markowitz, K. W. Kinzler and B. Vogelstein. 1996. *MAD*-related genes in the human. *Nature Genet.* **13**: 347-349.

Thiagalingam, S., C. Lengauer, F. S. Leach, M. Schutte, S. A. Hahn, J. Overhauser, J. K. V. Willson, S. Markowitz, S. R. Hamilton, S. E. Kern, K. W. Kinzler and B. Vogelstein. 1996. Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers. *Nature Genet.* **13**: 343-346.

Thiagalingam, S., N. A. Lisitsyn, M. Hamaguchi, M. H. Wigler, J. K. V. Willson, S. D. Markowitz, F.

S. Leach, K. W. Kinzler and B. Vogelstein. 1996. Evaluation of *FHIT* gene in colorectal cancers. *Cancer Res.* **56**: 2936-2939.

Thiagalingam, S., K. W. Kinzler and B. Vogelstein. 1995. *PAK1*, a gene that can regulate p53 activity in yeast. *Proc. Natl. Acad. Sci.* USA. **92**: 6062-6066.

Tokino, T., **S. Thiagalingam**, W. S. El-Deiry, T. Waldman, K. W.Kinzler, and B. Vogelstein.1994. *p53*-tagged sites from human genomic DNA. *Human Mol. Genet.* **3**: 1537-1542.

Pietenpol, J. A., T. Tokino, **S. Thiagalingam**, W. S. El-Deiry, K. W. Kinzler, and B. Vogelstein. 1994. Sequence-specific transcriptional activation is essential for growth suppression by p53. *Proc. Natl. Acad. Sci.* **91**:1998-2002.

Grossman, L., B. Ahn, L. Claassen, E. Hildebrand, S. Mazur, E. Oh, T. Seeley, **S. Thiagalingam**, and J. T. Wang. 1994. Nucleotide excision-repair - A tracking mechanism in search of damage. *J. Cell. Biochem.* Supplement **18C**: p129.

Oliner, J. D., J. A. Pietenpol, **S. Thiagalingam**, K. W. Kinzler, and Vogelstein, B. 1993. Oncoprotein MDM2 conceals the activation domain of p53. *Nature* **362**: 857-862

Grossman, L., and **S. Thiagalingam.** 1993. Nucleotide excision repair, a tracking mechanism in search of damage. *J. Biol. Chem.* **268**: 16871-16874.

Grossman, L., B. Ahn, L. Claassen, E. Hildebrand, S. Mazur, E. Oh, T. Seeley, **S. Thiagalingam**, and J. T. Wang. 1993. Nucleotide Excision Repair - A study in vectorial movement. *J. Cell. Biochem.* Supplement **17C**: p148.

Thiagalingam, S., and L. Grossman. 1993. The multiple roles for ATP in the *Escherichia coli* UvrABC catalyzed incision reaction. *J. Biol. Chem.* **268**: 18382-18389.

Thiagalingam, S., and T. Yang. 1993. Purification and characterization of NADH dehydrogenase from *Bacillus megaterium. Can. J. Microbiol.* **39**: 826-833.

Kern, S. E., J. A. Pietenpol, **S. Thiagalingam**, A. Seymour, K. W. Kinzler and B. Vogelstein. 1992. Oncogenic forms of p53 inhibit p53-regulated gene expression. *Science* **256**: 827-830.

Thiagalingam, S., and L. Grossman. 1991. Both ATPase sites of *Escherichia coli* UvrA have functional roles in nucleotide excision repair. *J. Biol. Chem.* **268**: 18382-18389.

Thiagalingam, S., and L. Grossman. 1991. Functional roles of the ATPase sites of *Escherichia coli* UvrAB in nucleotide excision repair. *J. Cell. Biochem.* Supplement **15D**: p110.

Oh, E. Y., L. Claassen, **S. Thiagalingam**, S. Mazur and L. Grossman. 1989. ATPase activity of UvrA and UvrAB protein complexes of the *Escherichia coli* UvrABC endonuclease. *Nucleic Acids Res.* **17**: 4145-4159.

Thiagalingam, S., Oh, E. Y., Mazur, S., and L. Grossman. 1988. Allosteric interactions between nucleotide and DNA binding sites in the *E. coli* UvrA protein. *J. Cell. Biochem.* Supplement **12A**: p278.

Thiagalingam, S., and T. Yang. 1986. NADH dehydrogenase and NADH oxidase in *Bacillus megaterium. Curr. Microbiol.* **14**: 217-220.

APPENDIX I -Reprints

1. Cheng, K-h., J. F. Ponte, and S. Thiagalingam. 2003. Elucidation of *Smad8* inactivation in cancer using Targeted Expressed Gene Display. *Manuscript submitted.*

2. Thiagalingam, S., K-h.Cheng, H. J. Lee, N. Mineva, and J. F. Ponte. 2003. Histone deacetylases: Unique players in shaping the epigenetic histone code, *Annal. New York Acad. Sci.* 983: 86-100.

Elucidation of epigenetic inactivation of *Smad8* in cancer using Targeted Expressed Gene Display

Kuang-hung Cheng[1,2], Jose F. Ponte[1] and Sam Thiagalingam[1-3]

Departments of [1]Medicine (Genetics and Molecular Medicine Programs and Pulmonary Center), [2]Pathology & Laboratory Medicine and [3]Genetics & Genomics, Boston University School of Medicine, 715 Albany Street, Boston, Massachusetts 02118

Correspondence should be addressed to S.T. (email- samthia@bu.edu)

Key words: TEGD, TGFβ, BMP, Smads, Smad8, DNA methylation, cancer, signature banding pattern (SBP).

Running title: Elucidation of *Smad8* silencing in cancer using TEGD.

(Submitted for Review).

ABSTRACT

To address the challenge of identifying related members of a large family of genes, their variants and their patterns of expression, we have developed a novel technique known as Targeted Expressed Gene Display (TEGD). Here we demonstrate the general application of this technique by analyzing the *Smad* genes, and report that the loss of *Smad8* expression is associated with multiple types of cancers, including 31% of both breast and colon cancers. Epigenetic silencing of *Smad8* expression by DNA hypermethylation in cancers directly correlates with loss of *Smad8* expression. The *Smad8* alteration in a third of breast and colon cancers makes it a significant novel tumor marker as well as a potential therapeutic target. The utility of TEGD as demonstrated by the analysis of *Smad* genes suggests that it is an efficient tool for the initial discovery of alterations in expressed genes within highly homologous gene families.

INTRODUCTION

Methods such as RT-PCR, cDNA subtraction, differential display (DD), representational difference analysis (RDA), serial analysis of gene expression (SAGE) and microarrays have been widely used in the identification of novel transcripts as well as in the assessment of their levels of expression in development, various cellular processes and diseases including cancer. Despite the usefulness of these techniques in the overall assessment of genes that are highly divergent at the DNA sequence, accurate and high throughput evaluation and discovery of related members of a gene family have remained a challenge. These methods in general have been unable to discriminate between different members of the gene families with consistency because of the inherent redundancy in DNA sequence among these unique genes and transcripts. A novel method described here, targeted expressed gene display (TEGD), validated using the *Smad* family of genes as the prototype, enables one to overcome this dilemma when gene family members contain at least two regions of homology separated by a divergent region of variable length.

The discovery of the Smad family of signal transducer proteins as mediators of TGFβ (transforming growth factor-beta) signaling from the cell membrane to the nucleus has revolutionized the understanding of the molecular basis of the signaling and inactivation of TGFβ/ BMP pathways in cancer (1). To date, eight human homologues of the *Smad* genes have been identified and are classified into three distinct classes based on their structures and biological functions (1,2). The first category consists of pathway-restricted or receptor-regulated Smads (R-Smads): Smad1, Smad5

and Smad8, which are involved in BMP signaling and Smad2 and Smad3 which are TGFβ /activin pathway restricted. These Smads are activated directly *via* phosphorylation by RI receptors following the formation of a complex consisting of the ligand bound heteromeric RI/ RII receptors. Phosphorylated R-Smads interact with the second class of Smads known as the common mediator Smad (Co-Smad) to form a heteromeric complex (3). Smad4 is the only member of this class of Smads known in mammals. The third class of Smads includes Smad6 and Smad7 which were identified as anti-Smads or inhibitory Smads (I-Smad) due to their ability to act as inhibitors of the signaling pathway (4-6).

Since the signaling pathways mediated by the members of the TGFβ family are implicated in a number of biological processes including cell differentiation, cell proliferation, determination of cell fate during embryogenesis, cell adhesion, cell death, angiogenesis, metastasis and immunosuppression, it is conceivable that genetic or epigenetic anomalies leading to altered expression patterns of various Smad molecules could contribute to different aspects of neoplastic progression (2, 7-10). Although there has been significant progress in elucidating the association between genetic alterations in the *Smad4* gene and cancer, the nature of defects involving the other Smads has been elusive (11-16). The apparent lack of genetic alterations in the majority of *Smad* genes analyzed thus far in cancer provides compelling support for the potential role of epigenetic alterations, whereby abnormalities in signaling could occur at the level of regulation of gene expression or processing of the transcripts (17-19). Our analysis of the *Smad* genes provides evidence for the exploitation of the novel TEGD method

described in this article in the initial determination of the mode of inactivation of the *Smad* genes in cancer. Thus, we predict that the effective utilization of the method described here will find wide use not only in the discovery of novel members of a family of genes and splice variants of a specific gene, but also for the simultaneous analysis of the transcript levels of individual genes or their spliced variants in various diseases and during development.

MATERIALS AND METHODS

Cell culture, RNA isolation and cDNA synthesis.

Cancer cell lines were purchased from ATCC or Coriell Cell Repository and culture conditions were followed as suggested by the provider. Tumor samples, some of the cell lines and their derivatives or nucleic acids isolated from the samples used in this study were obtained from Subra Kugathasan (Medical College of Wisconsin), Peter Thomas (Boston University School of Medicine), Douglas Faller (Boston University School of Medicine), Ramon Parsons (Columbia University) and Kornelia Polyak (Dana Farber Cancer Institute). RNA isolation and cDNA synthesis from the cell lines and tumor samples were carried out using previously described procedures (20).

Smad genes degenerate RT-PCR.

Based on the amino acid sequences of the human Smads 1-8, regions that are identical and conserved (MH1 and MH2) among the Smads were mapped out (1, 2). The residues targeted for the primer design were localized to the MH1 and MH2 domains, and the intervening linker regions were highly divergent enabling the generation of PCR products that are of unique size

corresponding to specific *Smad* homolog(s). The forward and reverse primers were designed based on the maintenance of codon degeneracy and the representation of the various amino acids at a given position among the known Smad family members as determined from the sequence alignment of the various homologs. All primers were obtained from Integrated DNA Technologies, Coralville, IA.

The *Smad* family specific degenerate primers used for TEGD are as follows: SmadXF2(5' primer) -TNTKBMGVTGGCCNGAYYTBM; SmadXR1(3' primer) - CCAVCCYTTSRCRAARCTBAT (Codes for mixing of bases to generate degeneracy: R=A,G; Y=C,T; M=A,C; K=G,T; S=C,G; W=A,T; H=A,C,T; V=A,C,G; D=A,G,T; N=A,C,G,T).

A 20 μl PCR reaction mixture contained 67 mM Tris-HCl, pH 8.8, 16.6 mM ammonium sulfate, 6.7 mM magnesium chloride, 1 mM β-mercaptoethanol, 6% dimethyl sulphoxide, 100 μM each of dATP, dGTP, dCTP and dTTP, radiactive dCTP (0.25 μl of α^{32}P- dCTP (10 μCi/μl), Amersham) for labeling, 20 μM each of the primers, 50 ng of cDNA template and 2.5 Units of Platinum Taq (Invitrogen). An initial denaturation at 94^{0}C for 2 minutes was followed by 30 cycles, each carried out at 94^{0}C for 30 seconds, 57^{0}C for 1 minute, and 70^{0}C for 1 minute and 20 seconds; and one final extension cycle at 70^{0}C for 10 minutes to facilitate TA cloning into pCR2.1 (Invitrogen).

TEGD gel electrophoresis and recovery of DNA bands.

The samples from the degenerate RT-PCR of the *Smad* genes were loaded onto a 4.5% denaturing polyacrylamide gel after a 2 minute

denaturation step at 95^{0}C. Electrophoresis was performed in a Genomyx LR analyzer (Beckman Coulter) for 4.5 hrs at 80 Watts with constant power (voltage not to exceed 2500 volts). The gel was dried and autoradiography performed on the gel. DNA bands of interest on the gel were oriented using the autoradiogram, cut out of the gel and isolated by soaking the gel slice in 1XTE buffer, freezing at -80°C for 30 minutes, heating at 60°C for 5 minutes and spinning at high speed to separate gel fragments from the aqueous phase containing DNA. The DNA fragments were ethanol precipitated and isolated using conventional methods and TA cloned (Invitrogen) for sequencing.

DNA sequencing.

DNA sequence analysis was performed using the Genomyx LR analyzer (Beckman Coulter). The cycle sequencing procedure used in these studies utilized ^{33}P ddNTPs (Amersham) along with the ThermoSequanase kit (USB, Cincinnati, OH) as previously described (20).

Genomic DNA Isolation.

Genomic DNA from cell lines and tumors were isolated using the DNeasy Tissue Kit (QIAGEN) according to the manufacturer's instructions.

Homozygous deletion analysis of Smad8.

Radiolabeled microsatellite markers, D13S927 and D13S928, that are localized at the beginning and end, respectively, of the *Smad8* gene in its genomic contig and gene specific primers encompassing the first (*Smad8* EX-

1F: 5'-GAAACATGTGAGGAACAGCAGC-3' and *Smad8* EX-1R: 5'-CGAGACAGCGGCTGCAGCAGCG-3') and the second exons (*Smad8* EX-2F: 5'-GCCTGGTTCTGTTGCTCAGGCTG-3' and *Smad8* EX-2R: 5'-GTGTTCCTGTGGCATTCAGGC-3') of the *Smad8* gene were used in PCR amplifications and gel electrophoretic analysis to determine deletion of this genomic region (12).

Analysis of gene expression using semi-quantitative RT-PCR.

Total RNA prepared from samples was used for cDNA synthesis and PCR amplification was done essentially as previously described (20). The gene specific primer pairs used in the analysis of the indicated specific *Smad* genes and the β-actin gene used for standardization to normalize the abundance of the various transcripts analyzed are as follows:

Smad1-F: 5'-CCACTGGAATGCTGTGAGTTTCC-3'

Smad1-R: 5'-GTAAGCTCATAGACTGTCTCAAATCC-3'

Smad2-F: 5'-GGTAAGAACATGTCCATCTTGCC-3' (20)

Smad2-R: 5'-CATGGGACTTGATTGGTGAAGC-3' (20)

Smad3-F: 5'-CGGGCCATGGAGCTGTGTGAGTTCG-3'

Smad3-R: 5'-CGGGTCAACTGGTAGACAGCCTC-3'

Smad4-F: 5'-GGACAATATGTCTATTACGAATAC-3' (20)

Smad4-R: 5'-TTTATAAACAGGATTGTATTTTGTAGTCC-3' (20)

Smad5-F: 5'-GTATCAACCCATACCACTATAAGAG-3'

Smad5-R: 5'-CAGAGGGGAGCCCATCTGAGTAAG-3'

Smad7-F: 5'-GGTGCGAGGTGCCAAATGTCACC-3'

Smad7-R: 5'-GATGAACTGGCGGGTGTAGCAC-3'

Smad8-F: 5'-CTCTTATGCACTCCACCACCCCCATC-3'

Smad8-R: 5'-CTTAAGACATGACTGTTAAGACACTG-3'

β-Actin-F: 5'-ACACTGTGCCCATCTACGAGG-3'

β-Actin-R: 5'-AGGGGCCGGACTCGTCATACT-3'

The relative abundance of the various *Smad* gene-specific PCR products was normalized to β-Actin or other unaffected Smads by comparative abundance of the products using densitometry.

Processing of genomic DNA for the evaluation of methylation status.

For bisulfite sequencing and the MSP assay, genomic DNA was isolated from cell lines and primary tumors using the QIAGEN DNeasy Tissue Kit. Genomic DNA was subjected to a deamination reaction by incubation with sodium bisulfite essentially as previously described (21). In brief, 0.5 to 2 μg genomic DNA was denatured with 2 M NaOH for 10 min, followed by bisulfite modification by treatment with freshly prepared 10 mM hydroquinone and 3 M sodium bisulfite, pH 5.0 (Sigma), which converts unmethylated cytosines to uracil but does not change methylated cytosines. Each reaction was overlaid with mineral oil and incubated at 50°C for 16-20 hours. After treatment, the modified DNA was purified using a Wizard DNA purification kit (Promega, Madison, Wisconsin), followed by desulfonation by treating with 3 M NaOH. The ethanol precipitated purified DNA pellet was dissolved in 30 μl of distilled water.

Bisulfite sequencing.

The intron 1 region of the *Smad8* gene containing CpG islands was first PCR amplified from bisulfite modified DNA (50-100 ng) using gene specific primers (5'-GAAATATGTGAGGAATAGTAGTTTAG-3' and 5'-CCACTCATCCCTCCCCCACCCAAATC-3') and the product was gel purified. Genomic sequencing of the *Smad8* gene-specific PCR product was accomplished by using the DNA sequencing primer, 5'-GTAAGTAGGGTTTTTTGGT-3', along with ^{33}P ddNTPs and the ThermoSequanase kit (USB, Cincinnati, OH) as previously described (20).

Methylation-Specic PCR (MSP).

The methylation status of the *Smad8* promoter region was also analyzed by MSP with the use of primers designed for the amplification of defined CpG islands containing DNA sequences of either unmethylated or methylated DNA (21). Sequences of the forward (F) and reverse (R) MSP primers to distinguish the methylated (M) and unmethylated (U) genomic DNA used in this study were as follows: 5'-GATGTGAGGTGATTTATGTAGT-3' (Smad8U-F) and 5'-CACAACAACCTACAACTCAATTCCCT-3' (Smad8U-R), and 5-GACGCGAGGCGATTTACG-3' (Smad8M-F) and 5'-CGACCACGTACGCGAAAACTCGCG-3' (Smad8M-R). PCR conditions were as follows: 94°C for 2 min, 35 cycles of 94°C for 30 sec, 58°C for 30 sec, 70°C for 40 sec, followed by a final extension at 70°C for 10 min. A 10 μL sample of each PCR product was mixed with 1 X loading buffer and analyzed by electrophoresis on a nondenaturing 8% polyacrylamide gel and visualized by staining with ethidium bromide.

5'-Aza-2' deoxycytidine and TSA treatment.

HTB129, MDAMB468, MDAMB231, CaCo2, H441, CCL230 and HT29 cells were incubated in culture medium with and without 5'-Aza-2' deoxycytidine (Sigma) at a concentration of 1-5 μM for 7 days or with 300 nM trichostatin A (TSA) for 24hrs. To assess the effect of a combination of 5'-Aza-2' deoxycytidine and TSA, cells were exposed sequentially for 7 days to 5'-Aza-2' deoxycytidine and then to TSA for an additional 24 hrs. Total RNA was isolated and *Smad8* expression was determined by RT-PCR using the primers Smad8-1F: 5'-CAGCTCAGCCTCCTGGCCAAG-3' and Smad8-1R: 5'-GAGGAAGCCTGGAATGTCTC-3'.

RESULTS

TEGD and signature banding pattern of the *Smad*s.

Members of the *Smad* family of genes have highly homologous amino acid sequences at their N- and C- terminal regions (MH1 and MH2-domains, respectively), which are separated by a highly divergent linker region rich in proline, serine and threonine (1, 2). These regions may have arisen from divergence due to functional specificities from an ancesteral unit of activity that has maintained some degree of evolutionary conservation at the level of the protein. The examination of the MH domains from various Smad genes indicated that there is identity and conservation among amino acid residues at defined positions, which is consistent with critical structural features required for the function of these proteins. The sequence conservation at the amino acid level is also reflected at the DNA sequence level. Despite, their similarity at the level of the genetic code, the Smad proteins are involved in a wide

array of cellular functions as they not only play roles as mediators, inhibitors and transcription factors of the Smad signaling pathways but also mediate signaling in response to a diverse but related cytokines (TGFβ family). Even though the delineation of the alterations in Smads is essential for the comprehension of the molecular basis of various defective processes, the analysis of defects in individual members in this type of family of genes poses a formidable task for efficient detection in a high throughput platform. Success in identifying alterations in *Smad* genes could be expected to provide critical information necessary for deciphering the molecular basis of their functions. The fact that the *Smad* genes contain two distinct highly conserved regions separated by a highly variable intervening linker region allowed us to develop a novel screening strategy to simultaneously analyze all the known members of this family (Figure 1A).

We have designed degenerate oligonucleotide primers corresponding to the conserved regions of the Smad family of genes based on the preservation of codon degeneracy and conserved amino acids at a given position among the known Smads for PCR amplification of the cDNA templates. PCR amplification in the presence of radiolabeled nucleotides, and the subsequent analysis of the products using a denaturing polyacrylamide gel electrophoresis revealed distinct bands on the gel (Figure 1 A & B). We recovered the distinct bands corresponding to the PCR products generated using Smad-specific degenerate primers and sequenced them using primers that are specific for the predicted *Smad* gene(s). The bands corresponding to the 1200, 960, 840, 680 and 570 base pairs (bp) PCR products were found to be identical to the cDNA sequences for *Smad4*, *Smad1* and *Smad5*, *Smad2*,

Smad3 and *Smad8*, and *Smad6* and *Smad7*, respectively, as predicted from their estimated sizes and sequences (1, 2; Figure 1B). These results suggested to us that once the signature banding pattern (SBP) of the targeted expression gene display is optimized and established, such as in this case with the *Smad* family of genes, repeat analysis of gene expression in tissues or other samples of unknown origin could be easily adopted for a routine high throughput analysis. Although we generated and analyzed radiolabeled PCR products in these initial studies, one could also achieve the same results using fluorescently or radioactively end-labeled primers for PCR amplification.

Validation of *Smad* expression patterns determined from TEGD.

We confirmed the presence or absence of *Smad* expression determined from TEGD using gene specific primers by semi-quantitative RT-PCR (Figure 1C). The expression patterns of the various Smads detected by TEGD remained consistent with semi-quantitative RT-PCR results. Most of the Smads were expressed in all the tissue types that we have analyzed, however, some Smad expression was lost in the liver and was decreased to barely detectable levels in the bone marrow and uterus (Figure 1C). These results indicated to us that TEGD could be used as a tool for initial diagnostic high throughput evaluations to determine *Smad* gene expression patterns simultaneously and with a high degree of efficiency. Thus, TEGD can be regarded as a highly improved alternate method that may substitute for the traditional multiplex PCR technique due to its increased level of sensitivity, ability to discriminate between genes that are closely related at their DNA sequence and the low level of cDNA template required for the analysis.

Differentially spliced variants of the *Smad*s.

TEGD also enabled us to identify the various differentially spliced forms of the *Smad2*, *Smad3*, *Smad5*, and *Smad8* genes (Figures 1B & C; data not shown). Alternatively spliced variants of *Smad2* with a deletion of exon3 (*Smad2*Δexon3), *Smad3* with deletions of both exons 3 and 7 (*Smad3*Δexon3 Δexon7), *Smad5* with a deletion of exon3 (*Smad5*Δexon3) and *Smad8* with deletions of either exon3 (*Smad8*Δexon3) or both exons 2 and 3 (*Smad8*Δexon2Δexon3) were detected in our analysis. Although one of these variants (*Smad2*Δexon3) has been previously reported, the existence of the others has been recorded for the first time in this study (22). However, our study did not verify the existence of two previously reported alternatively spliced forms with deletions at the 3' ends, potentially due to the placement of the TEGD primers inside the affected sequence of these alternatively spliced forms (23, 24). The encoded proteins of Smad2, Smad5 and Smad8 resulting from full-length and variant transcripts that have been described also exhibit differences in their biochemical properties (22-24). Despite these findings, the overall significance of the described and predicted novel spliced forms of *Smad*s reported both here and elsewhere in disease phenotypes including cancer requires further studies (22-24; UCSC genome browser: http://genome.ucsc.edu).

TEGD in the analysis of *Smad*s in cancer.

The ability to simultaneously probe multiple members of a gene family using TEGD prompted us to apply this technique to analyze differential

expression patterns of the various *Smad*s in cancer to validate its utility for diagnostic screening (Figure 2A). We were able to utilize the signature banding patterns established with the normal tissues to determine the retention or loss of specific DNA bands corresponding to the defined full-length and variant transcripts (Figures 1B and 2A). The TEGD analysis of the *Smad* genes in cancers lead us to conclude that there is a significant level of loss in the expression of *Smad3* and *Smad8* in colon cancer and of *Smad8* in breast cancer. These initial observations were further validated by analyzing the expression patterns of the *Smad8* gene more carefully using gene specific primers and semi quantitative RT-PCR (Figure 2B). These results further confirmed the TEGD data and provided the first clues to suggest that the *Smad8* gene is a critical target for loss of function due to down regulation of gene expression in 31% of breast and colon cancers (Table 1). The analysis to establish the significance of the loss of *Smad3* expression in colon cancer will be dealt with in greater detail elsewhere (Cheng and Thiagalingam, unpublished results). In conclusion, TEGD can be used as an initial dignostic tool in cancer and other diseases to simultaneously analyze differential expression patterns of genes that are closely related at the level of their nucleotide sequence.

Molecular mechanism for the silencing of *Smad8* expression.

From our analysis, loss of expression of the *Smad8* gene was estimated to occur in nearly a third of both breast and colon cancers, which are two of the leading causes of cancer deaths in women and in general, respectively (Figure 2B; Table1). Hence, we investigated potential mechanisms for the

loss of *Smad8* gene expression in cancer due to the high level of significance of this alteration with respect to the known tumor markers. We examined whether genetic alterations such as chromosomal deletions affecting the *Smad8* gene could lead to the loss of its expression by homozygous deletion analyses. We used microsatellite markers corresponding to the *Smad8* gene based on the genomic contig as well as by genomic PCR using primers that amplified the genomic region corresponding to the first two exons of the *Smad8* gene. These experiments indicated that gross genomic deletions are apparently not the major mechanism of *Smad8* inactivation in the affected cancers (data not shown). Therefore, we considered epigenetic silencing as an alternate mechanism for *Smad8* gene silencing.

The genomic sequence of the *Smad8* gene was inspected for the presence of CpG islands that may be the targets of DNA hypermethylation and associated chromatin modification effects for their involvement in the silencing of *Smad8* gene expression. Several CpG islands in the upstream promoter as well as in the first intronic region of the *Smad8* gene were tested as likely candidate regions that could be critical for differential DNA methylation patterns coinciding with the loss of *Smad8* expression (data not shown, Figure 3). DNA sequence analysis of the bisulfite treated genomic DNA revealed that CpG islands localized to nucleotides 3541028 to 35410583 (Chromosome 13q12-14 on the reverse strand between *Rb* and *BRCA2*; UCSC genome browser: http://genome.ucsc.edu) in the first intron of the *Smad8* gene are only methylated in cancers that exhibit loss of expression (Figure 3A & B). Methylation specificic PCR (MSP) was carried out using primers designed to these corresponding differentially methylated regions and

the results further confirmed that the *Smad8* gene is silenced in cancers due to DNA hypermethylation affecting CpG islands in the first intron of the *Smad8* gene (Figure 3C).

DNA hypermethylation and *Smad8* expression in cancer.

To directly determine the physiological significance, the role(s) of apparent epigenetic DNA methylation by itself or in combination with histone acetylation/ deacetylation on differential regulation of *Smad8* expression in cancers was examined. We chose six cell lines derived from breast, colon and lung cancers (HTB129, HT29, CaCo2, CCL253, MDAMB468 and H441) that exhibited loss of *Smad8* expression, and one cell line (MDAMB231) which retained *Smad8* expression as a control, and examined the effects of 5'-aza-2'-deoxycytidine (5Aza-dC; a DNA demethylating agent) and/or trichostatin A (an inhibitor of histone deacetylases) on *Smad8* expression. A substantial increase in *Smad8* expression was observed with 5Aza-dC treatment in all of the cell lines, which were previously determined to exhibit DNA hypermethylation-mediated gene silencing of *Smad8* (Figure 4A). Trichostatin A by itself caused only a slight increase in the levels of the transcript in two of the tested cell lines (CaCo2 and CCL253) but had no effect in the majority of the tested cell lines. However, there was a slight up regulation of Smad8 expression in the presence of both drugs (Figure 4A). MSP analysis of the target CpG islands in intron1 of the *Smad8* regulatory regions that were differentially methylated in affected and control cell lines revealed that demethylation due to 5Aza-dC treatment accompanies a corresponding increase in *Smad8* gene expression (Figure 4B). These

observations strongly support the notion that the loss of *Smad8* expression in cancers is primarily mediated by hypermethylation of *cis*-regulatory CpG islands of the gene.

DISCUSSION

The analysis of highly homologous members of a family of genes to detect and establish differential gene expression patterns as well as the genetic alterations responsible for cancer and other diseases with limited amounts of clinical sample has remained a formidable task. Efficient methods to simultaneously analyze the closely related yet functionally divergent genes belonging to families would not only be important in accurate diagnosis and prognostic evaluation of a disease but could also be exploited for the identification of pharmacogenetic targets to customize therapy. We propose that the TEGD technique described in this article can be effectively utilized to analyze families of genes that contain at least two stretches of conserved regions, which are separated by a divergent linker region of variable length. TEGD provides a distinct advantage over techniques such as differential display (DD), a comparable methodology, which has been adopted for the simultaneous analysis of multiple genes, due to the latter's inability to detect differential gene expression patterns of targeted and defined genes. Furthermore, even an improved version of DD designed to analyze related genes (e.g., kinases) still fell short of efficiently establishing distinct expression patterns of the related genes and failed to identify novel genes with different functional roles (25-28).

On the other hand, with TEGD, once a signature banding pattern of the

targeted expressed gene display is optimized and established with an array of different normal tissues, such as in this case with the Smad family of genes, repeat analysis of gene expression of samples of unknown origin could be easily carried out in a routine high throughput manner (Figures 1 and 2). We believe that the TEGD technique should sufficiently address the dilemma of efficient simultaneous expression pattern analysis of related genes with relatively minute amounts of samples in clinical and investigational research settings. The development of an algorithm to predict the suitability of the applications of TEGD based on the presence of two distinct homologous regions separated by an intervening variable region that would enable the establishment of signature banding patterns from the available sequences of already identified genes or ESTs is in progress. We believe that TEGD has the potential to advance the ability to probe gene families for genetic and epigenetic defects to a new level of sophistication and will find general use in the future. The application of the TEGD technique to simultaneously analyze multiple members of the *Smad* family of genes has not only validated the enormous advantage of the technique as an initial diagnostic tool but also illustrates an efficient way to identify novel genes that are closely related at the level of their nucleotide sequence, to identify splice variants of a gene as well as to detect their altered expression patterns.

Our survey of the various *Smad* genes using the novel TEGD technique described in this article enabled us to obtain the first clues in identifying the *Smad8* gene as an important target for loss of expression in multiple types of cancers, including nearly 31% of breast and colon cancers. This level of alteration is even more frequent than that of the *Smad4* gene, the most

frequent target for genetic inactivation of the known Smad signaling genes in colon cancer, and is also more frequent than the *HER/neu* gene amplification, the most celebrated tumor marker for breast cancer, which occurs in about 20%-30% of breast cancer cases (29).

Thus, the data presented in this article provides the first direct evidence that silencing of gene expression *via* DNA hypermethylation of the *Smad8* gene could be an important event in tumorigenesis of several cancers including one third of breast and colon cancers. It is interesting to note that *Smad8* is apparently the major target for loss of function among the *Smad* genes in breast cancer and is a R-Smad which becomes phosphorylated during BMP signaling events and modulates BMP-responsive genes including those that may affect bone homeostasis (30-34; Figure 5). Additionally Smad signaling events *via* the BMP cytokines are also implicated in other signaling events that regulate biological processes, including cell differentiation, proliferation, determination of cell fate during embryogenesis, cell adhesion, cell death, angiogenesis, metastasis and immunosuppression (1, 2; Figure 5). Although it is intriguing that metastasis to bone is often associated with advanced stage breast and other cancers, further studies would be required to understand whether metastatic breast cancer cells defective in Smad8 signaling could be responsible for causing an imbalance in normal bone homeostasis by enhancing osteoclastic bone resorption, leading to osteolytic lesions within the bone (35-38).

Additionally, despite the fact that inactivation of the *Smad2* and *Smad4* genes due to intrageneic mutations and homozygous deletions has been reported in nearly 20% of colorectal cancers, evidence for genetic or

epigenetic inactivation of other *Smad* gene targets at significant levels had remained elusive until this report (2, 20). The loss of expression of *Smad8* in nearly 31% of colon cancers is more significant than any other Smad alterations known to date. Determination of whether the affected cells play a critical role in tumorigenesis by a mechanism similar to that in breast cancer requires further study. Interestingly, the presence of germline mutations in the BMP receptor 1A in juvenile polyposis, which increase the risk of developing gastrointestinal cancers, suggests that inactivation of BMP signaling may play a critical role in colon cancer (39, 40). Despite the fact that the elucidation of BMP-mediated signaling pathways in which Smad8 is a critical mediator is still in its infancy, these studies clearly provide the incentive for further investigations that may help gain a better understanding of the effects of *Smad8* inactivation in cancer and could pave the way for the exploration of its potential utility in diagnosis, prognosis and designing of therapeutic modalities.

ACKNOWLEDGEMENTS

The authors would like to thank their collegues, Subra Kugathasan (Medical College of Wisconsin), Peter Thomas (Boston University School of Medicine), Douglas Faller (Boston University School of Medicine), Ramon Parsons (Columbia University) and Kornelia Polyak (Dana Farber Cancer Institute) for generously donating reagents and cell lines. This work was supported by grants from the Evans Medical Foundation and the Career Development Award to S.T. from the Department of Defense (DAMD 17-01-1-0160). J. F. P has received support from the NIH training grant

(T32HL07035). S.T. is a Dolphin Trust investigator supported by the Medical Foundation.

REFERENCES

1. Thiagalingam, S., Cheng, K-h., Foy, R. L., Lee, H. J., Chinnappan, D., and Ponte. J. F. TGF-β and its *Smad* connection to cancer. Current Genomics *3*, 449-476, 2002.

2. Massague, J. TGF-beta signal transduction. Annu. Rev. Biochem. *67*,753-791, 1998.

3. Abdollah, S., Macias-Silva, M., Tsukazaki, T., Hayashi, H., Attisano, L. and Wrana, J. L. TbetaRI phosphorylation of Smad2 on Ser465 and Ser467 is required for Smad2-Smad4 complex formation and signaling. J. Biol. Chem. *272*, 27678-27685, 1997.

4. Nakao, A., Afrakhte, M., Moren, A., Nakayama, T., Christian, J. L., Heuchel, R., Itoh, S., Kawabata, M, Heldin, N.E., Heldin, C.H. and ten Dijke. P. Identification of Smad7, a TGFbeta-inducible antagonist of TGF-beta signalling. Nature *389*, 631-635, 1997.

5. Imamura, T., Takase, M., Nishihara, A., Oeda, E., Hanai, J., Kawabata, M. and Miyazono, K. Smad6 inhibits signalling by the TGF-beta superfamily. Nature *389*, 622-626, 1997.

6. Hayashi, H., Abdollah, S., Qiu, Y., Cai, J., Xu, Y.Y., Grinnell, B. W., Richardson, M.A., Topper, J.N., Gimbrone, M. A. Jr., Wrana, J.L. and Falb, D. The MAD-related protein Smad7 associates with the TGFbeta receptor and functions as an antagonist of TGFbeta signaling. Cell *89*, 1165-1173, 1997.

7. Massague, J. and Wotton, D. Transcriptional control by the TGF-beta/Smad signaling system. EMBO J. *19*, 1745-1754, 2000.

8. de Caestecker, M.P., Piek, E. and Roberts, A.B. Role of transforming growth factor-beta signaling in cancer. J. Natl. Cancer Inst. *92*, 1388-1402, 2000.

9. Massague, J., Blain, S.W. and Lo, R.S. TGFbeta signaling in growth control, cancer, and heritable disorders. Cell *103*, 295-309, 2000.

10. Letterio, J.J. and Roberts, A.B. Regulation of immune responses by TGF-beta. Annu. Rev. Immunol. 16,137-161, 1998.

11. Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R.S., Zborowska, E., Kinzler, K.W., Vogestein, B. and Wilson,

J.K.V. Inactivation of the type II TGF-β receptor in colon cancer cells with microsatellite instability. Science *268*,1336-1338, 1995.

12. Thiagalingam, S., Lengauer, C., Leach, F. S., Schutte, M., Hahn, S. A., Overhauser, J., Willson, J.K.V., Markowitz, S., Hamilton, S.R., Kern, S.E., Kinzler, K.W. and Vogelstein, B. Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers. Nature Genet. *13*, 343-346, 1996.

13. Eppert, K., Scherer, S.W., Ozcelik, H., Pirone, R., Hoodless, P., Kim, H., Tsui, L.-C., Bapat, B., Gallinger, S., Andrulis, I. L., Thomsen, G. H., Wrana, J. L. and Attisano, L. MADR2 maps to 18q21 and encodes a TGF-β regulated MAD related protein that is mutated in colorectal carcinoma. Cell *86*, 543-552, 1996.

14. Riggins, J. G., Thiagalingam, S., Rozenblum, E., Weinstein, C. L., Kern, S. E., Hamilton, S. R., Willson, J.K.V., Markowitz, S., Kinzler, K.W. and Vogelstein, B. *MAD*-related genes in the human. Nature Genet. *13*, 347-349, 1996.

15. Schutte, M., Hruban, R.H., Hedrick, L., Cho, K.R., Nadasdy, G.M., Weinstein, C.L., Bova, G.S., Isaacs, W.B., Cairns, P., Nawroz, H., Sidransky, D., Casero, R.A. Jr., Meltzer, P.S., Hahn, S.A. and Kern S. E. DPC4 gene in various tumor types. Cancer Res. *56*, 2527-2530, 1996.

16. Riggins, J. G., Kinzler, K. W., Vogelstein, B. and Thiagalingam, S. Frequency of *Smad* gene mutations in human cancers. Cancer Res. *57*, 2578-2580, 1997.

17. Thiagalingam, S., Cheng, K-h., Lee, H. J., Mineva, N., Thiagalingam, A. and Ponte, J. F. Histone deacetylases: Unique players in shaping the epigenetic histone code, Annal. New York Acad. Sci. *983*, 84-100, 2003.

18. Baylin, S.B. Tying it all together: epigenetics, genetics, cell cycle, and cancer. Science *277*,1948-1949, 1997.

19. Jones, P. A. and Laird, P. W. Cancer epigenetics comes of age. Nat. Genet. *21*, 163-167, 1999.

20. Thiagalingam, S. Molecular detection of Smad4/Smad2 alterations in colorectal tumors: Colorectal Cancer Methods and Protocols. *In*: S. M. Powell (Ed) Methods in Molecular Medicine. *50*, pp149-165. New Jersey: Humana Press Inc., 2001.

21. Herman, J. G., Graff, J. R., Myohanen, S., Nelkin B. D. and Baylin, S. B. Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. Proc. Natl. Acad. Sci. USA, *93*, 9821-9826, 1996.

22. Yagi, K., Goto, D., Hamamoto, T., Takenoshita, S., Kato, M. and Miyazono, K. Alternatively spliced variant of Smad2 lacking exon 3.

Comparison with wild-type Smad2 and Smad3. J Biol Chem. *274*, 703-709, 1999.

23. Jiang, Y., Liang, H., Guo, W., Kottickal, L.V. and Nagarajan, L. Differential expression of a novel C-terminally truncated splice form of SMAD5 in hematopoietic stem cells and leukemia. Blood *95*, 3945-3950, 2000.

24. Nishita, M., Ueno, N. and Shibuya, H. Smad8B, a Smad8 splice variant lacking the SSXS site that inhibits Smad8-mediated signalling. Genes Cells *4*, 583-591, 1999.

25. Liang, P. A. decade of differential display. Biotechniques *33*, 338-344, 346, 2002.

26. Okada, M., Yamaga, S., Yasuda, S., Weissman, S.M. and Yasukochi, Y. Differential display of protein tyrosine kinase cDNAs from human fetal and adult brains. Biotechniques *32*, 856, 858, 860, 863-5, 2002.

27. Bosch, I., Melichar, H. and Pardee, A.B. Identification of differentially expressed genes from limited amounts of RNA. Nucleic Acids Res. 28, e27, 2000.

28. Ranamukhaarachchi, D.G., Rajeevan, M.S., Vernon, S.D. and Unger, E.R. Modifying differential display polymerase chain reaction to detect relative changes in gene expression profiles. Anal. Biochem. *306*, 343-346, 2002.

29. Zujewski, J.A. "Build Quality in"-HER2 Testing in the Real World. J.Natl. Cancer Inst. *94*, 788-789, 2002.

30. Hogan, B.L.M. Bone morphogenetic proteins: multifunctional regulators of vertebrate development. Genes Dev. *10*,1580-1594, 1996.

31. Hogan, B.L.M. Morphogenesis. Cell *96*, 225-233, 1999.

32. Massague, J. and Chen, Y. G. Controlling TGF-beta signaling. Genes Dev. *14*, 627-644, 2000.

33. Attisano, L. and Wrana, J. L. Smads as transcriptional co-modulators. Curr Opin Cell Biol. 12, 235-243, 2000.

34. Miyazono, K., Kusanagi, K. and Inoue, H. Divergence and convergence of TGF-beta/BMP signaling. J.Cell Physiol. *187*, 265-276, 2001.

35. Coleman, R.E., Smith, P. and Rubens, R.D. (1998) Clinical course and prognostic factors following bone recurrence from breast cancer. Br. J. Cancer. *77*, 336-340.

36. Coleman, R.E. and Rubens, R.D. The clinical course of bone metastases from breast cancer. Br J Cancer. *55*, 61-66, 1987

37. Reddi, A.H., Roodman, D., Freeman, C. and Mohla, S. Mechanisms of tumor metastasis to the bone: challenges and opportunities. J. Bone Miner Res. *18*, 190-194, 2003.

38. Reddi, A.H. Role of morphogenetic proteins in skeletal tissue engineering and regeneration. Nat. Biotechnol. *16*, 247-252, 1998.

39. Howe, J.R., Howe, J.R, Bair, J.L., Sayed, M.G., Anderson, M.E., Mitros, F.A., Petersen, G.M., Velculescu, V.E., Traverso, G. and Vogelstein, B. Germline mutations of the gene encoding bone morphogenetic protein receptor 1A in juvenile polyposis. Nat Genet. *28*, 184-187, 2001.

40. Sayed, M.G., Ahmed, A.F., Ringold, J.R., Anderson, M.E., Bair, J.L., Mitros, F. A., Lynch, H.T., Tinley, S.T., Petersen, G.M., Giardiello, F.M., Vogelstein, B and Howe, J.R. Germline SMAD4 or BMPR1A mutations and phenotype of juvenile polyposis. Ann. Surg. Oncol. *9*, 901-906, 2002.

Table:

Table 1. Altered expression of Smad genes in cancers.

Cancer	Total # of samples	Samples with loss of *Smad8* expression (%)
Breast	35	11/35 (31)
Colon	41	13/41 (31)
Esophagus	4	0/4 (0)
Head & Neck	4	2/4 (50)
Lung	19	1/19 (5)
Pancreas	3	2/3 (65)
Prostate	4	3/4 (75)
Ovary	2	1/2 (50)
Stomach	4	2/4 (50)

Figure legends:

Figure 1. Targeted expressed gene display (TEGD) and tissue-wide expression of *Smad* genes.

A. Schematic representation of TEGD for the Smad family of genes.
MH1 and MH2 indicate highly homologous regions in the amino acid as well as DNA sequence among the various *Smad* gene family members. The forward and reverse primers for PCR amplification of the cDNA were designed in the conserved regions as indicated. The radiolabeled PCR products were analyzed by denaturing acrylamide gel electrophoresis. A typical signature banding pattern (SBP) of the various Smads is indicated in the lower panel.

B.TEGD analysis of the *Smad* family of genes in various tissue types.
PCR products for *Smad*s using degenerate primers were analyzed by TEGD. Lanes 1-17 correspond to PCR products generated using cDNA templates from brain, lung, stomach, heart, liver, spleen, kidney, colon, bone marrow, small intestine, trachea, prostate, uterus, thymus, testis, skeletal muscle, and mammary gland, respectively. The lines on the right hand panel point to distinct PCR products. The approximate size of PCR products in base pairs (bp) is indicated on the left panel. The positions of various Smad genes and their variants as identified from sequence analysis are indicated on the right panel.

C. RT-PCR analysis of *Smad* genes.
Semi-quantitative RT-PCR analysis of the indicated Smad genes was carried out as described under materials and methods. The cDNA template was derived from total RNA from normal tissues of brain, lung, heart, liver, bone marrow, kidney, spleen, thymus, prostate, testis, uterus, small intestine, mammary gland, skeletal muscle, stomach and colon, lanes 1-16, respectively.

Figure 2. Analysis of *Smad* expression in cancer.

A. TEGD analysis of *Smad* genes in cancer.
PCR products of *Smad*s generated using degenerate primers as described under Figure 1 were obtained from differerent cancers and analyzed by TEGD. The cDNA templates used in reactions analyzed on lanes NC, NB & NS are from normal cells from colon, breast and stomach tissues; C1-7, B1-7 and S1-4 are from colon, breast and gastric cancers, respectively. The arrows point to distinct PCR products that were absent compared to the normal control. The positions of various Smad genes and their variants as identified from sequence analysis are indicated on the right panel.

B. *Smad8* expression in cancer cell lines and tumors.
Total RNA prepared from cell lines and tumors from the lung, breast and colon cancers were analyzed by RT-PCR (Lanes 1-14). Lane 1 in each of the different RT-PCR panels corresponds to the normal sample of the indicated tissue type. *Smad8*α, *Smad8*β and *Smad8*γ are three of the major

differentially spliced forms of *Smad8*, which correspond to transcripts that are full-length, that exhibit deletion of exon 2, and that exhibit deletions of exons 2 and 3, respectively. Analysis of the *β-Actin* gene is used for normalization and quantitation of the expression of *Smad8*.

Figure 3. Epigenetic gene silencing of the *Smad8* due to altered DNA methylation.

A. Schematic drawing of the landscape of CpG island methylation patterns in the region of genomic DNA from upstream of exon 1 through exon 2 of the *Smad8* gene. Boxes denote exons. The flag represents the ATG corresponding to the first methionine of the predicted peptide. Vertical lines indicate CpG islands in the DNA sequence. Open circles represent unmethylated cytosines whereas filled circles represent methylated cytosines as determined by bisulfite sequencing. The circles above the horizontal line indicate the methylation pattern observed in the CpG islands of the cell lines that express *Smad8*. The circles below the line indicate the methylation pattern of the CpG islands from samples, which lacked *Smad8* expression. The nucleotide sequence of the DNA within the dotted lines is shown with the asterisks (*) indicating CpG islands.

B. Bisulfite sequence analysis of the indicated CpG islands of intron 1 of the *Smad8* gene in the cell lines that are either proficient (++) or deficient (-) in *Smad8* expression. Cell lines proficient for *Smad8* expression have no bands in the C lane indicative of conversion of unmethylated cytosines to uracil upon bisulfite treatment.

C. MSP (Methylation Specific PCR) analysis of the various cancers that have lost or retained *Smad8* expression. The MSP products in lanes U and lanes M indicate the presence of unmethylated and methylated templates, respectively. Placental DNA (PDNA) and *in vitro* methylated DNA (IVM) serve as negative and positive controls.

Figure 4. The effects of DNA demethylation and inhibition of histone deacetylases on *SMAD8* gene expression.

The indicated cell lines were treated with 5-AZA-dC for 7 days or with TSA for 24hrs. To assess the effect of both 5-AZA-dC and TSA simultaneously, cells were exposed sequentially for 7 days to 5-AZA-dC and subsequently to TSA for an additional 24 hrs. Total RNA and genomic DNA were isolated and *Smad8* expression and DNA hypermethylation were determined by (A) RT-PCR and (B) MSP analysis, respectively. MDAMB231 cells were used as the positive control.

Figure 5. A model for the Smad8 connection to cancer.

BMP signaling is initiated by the association between BMPs and type I (RI) and type II (RII) heteromeric receptors which follows phosphorylation of the type I receptor (RI) kinase that in turn phosphorylates the receptor-regulated Smads (R-Smad), such as Smad8 and initiates the signaling events. The

phosphorylated Smad8 forms a heteromeric complex with the common-mediator Smad (Co-Smad), Smad4, and is translocated into the nucleus. In the nucleus, the Smad8/Smad4 hetero-oligomer either by itself or by associating with heterologous Smad-interacting DNA binding proteins (SIDBP) or other cofactors, could mediate specific transcriptional activation or repression responses. The inhibitory Smads (I-Smad) such as Smad6 and Smad7 are able to compete with the R-Smads by stably binding the RI kinase or by preventing association of R-Smads with the Co-Smad, effectively blocking the signaling cascade. There are numerous other signaling pathways such as the Ras-MEK pathway that could also modulate the end effects by establishing cross talk among the different pathway members. BMP signaling is implicated in tumor suppression, bone homeostasis, angiogenesis and metastasis.

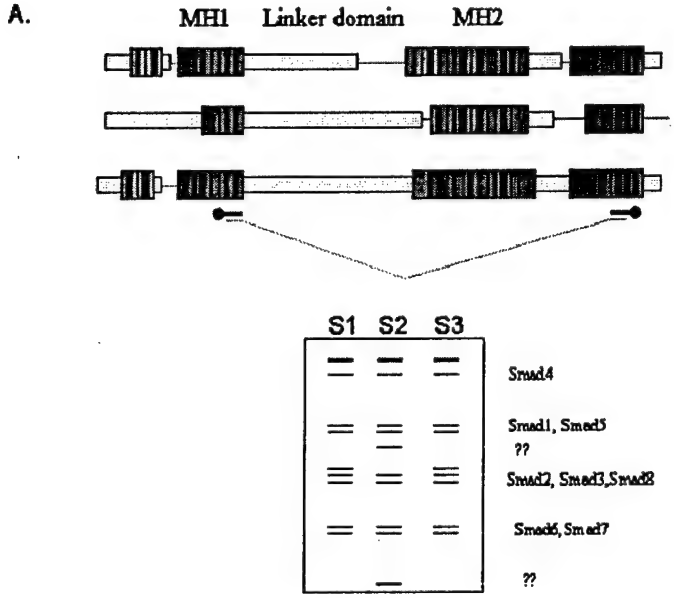

Figure 1A. Targeted Expressed Gene Display (TEGD) and tissue wide expression of Smads.

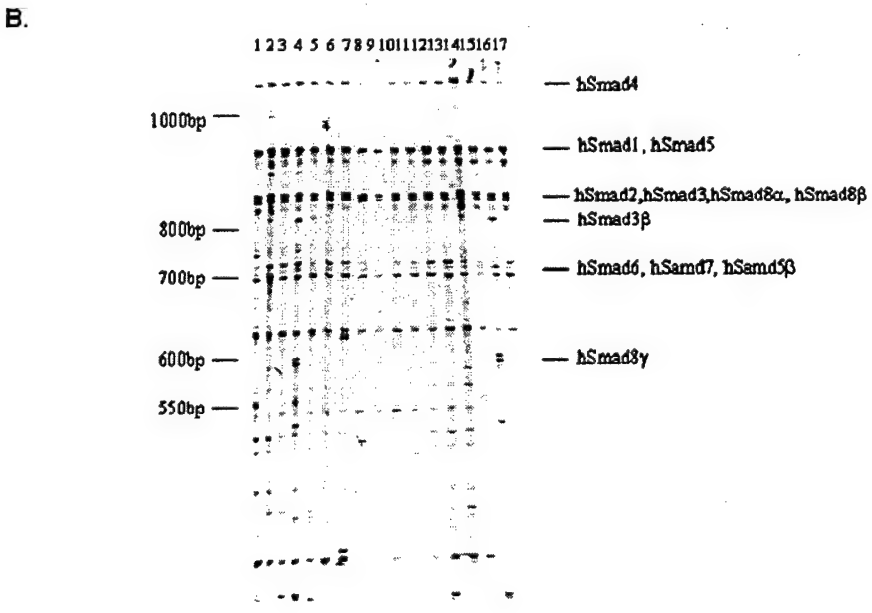

Figure 1B. Targeted Expressed Gene Display (TEGD) and tissue-wide expression of Smads.

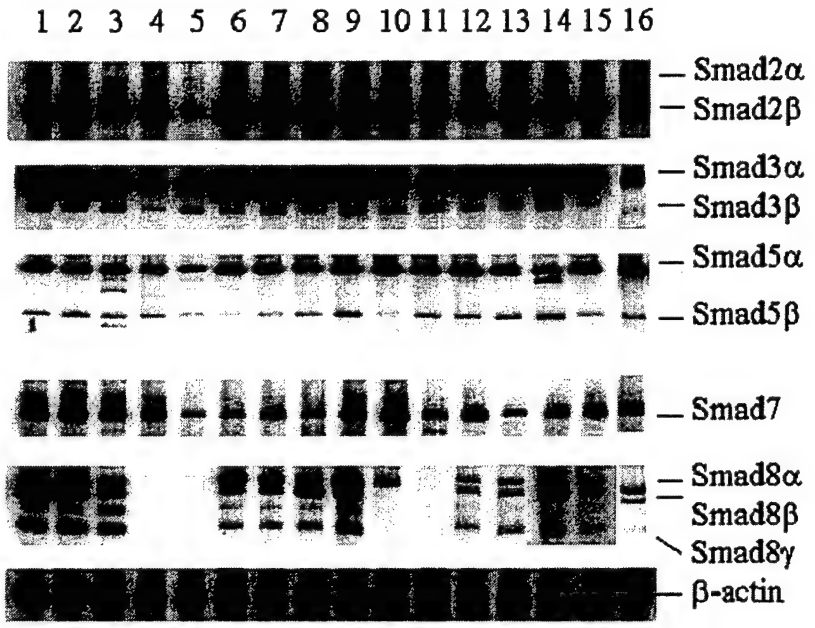

Figure 1C.Target Expressed Gene Display (TEGD) and tissue-wide expression of *Smads*.

A

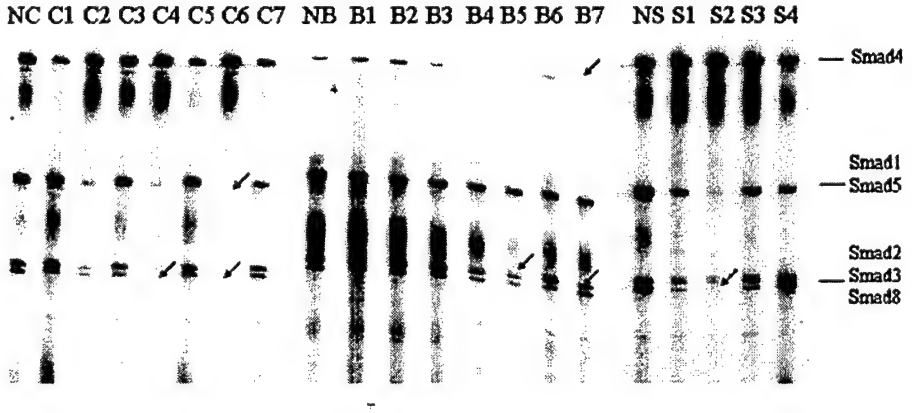

Figure 2A. Analysis of *Smad* genes expression in cancer.

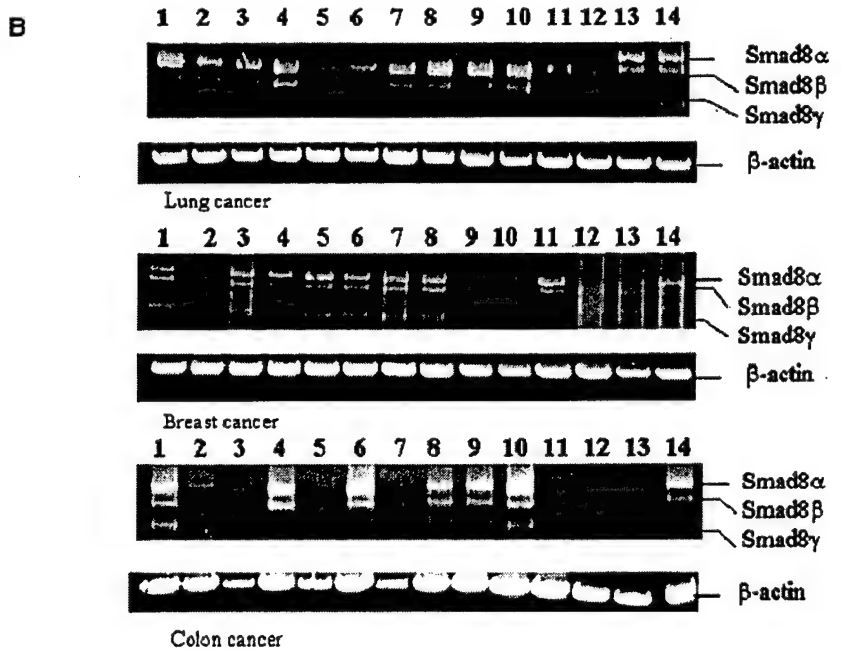

Figure 2B. Analysis of Smad genes expression in cancers.

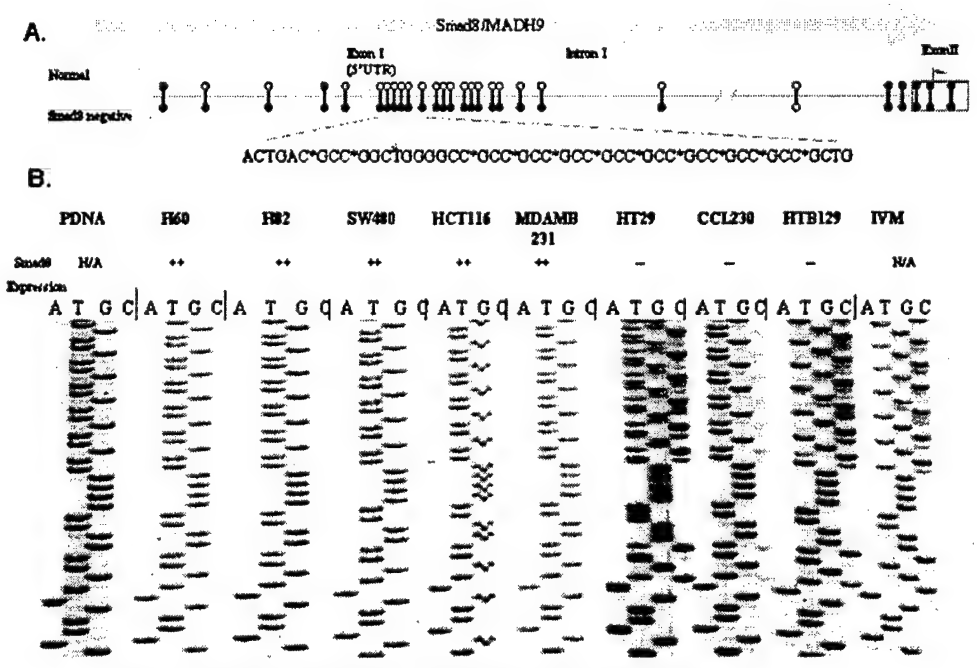

Figure 3. Epigenetic gene silencing of the *Smad8* due to altered DNA methylation.

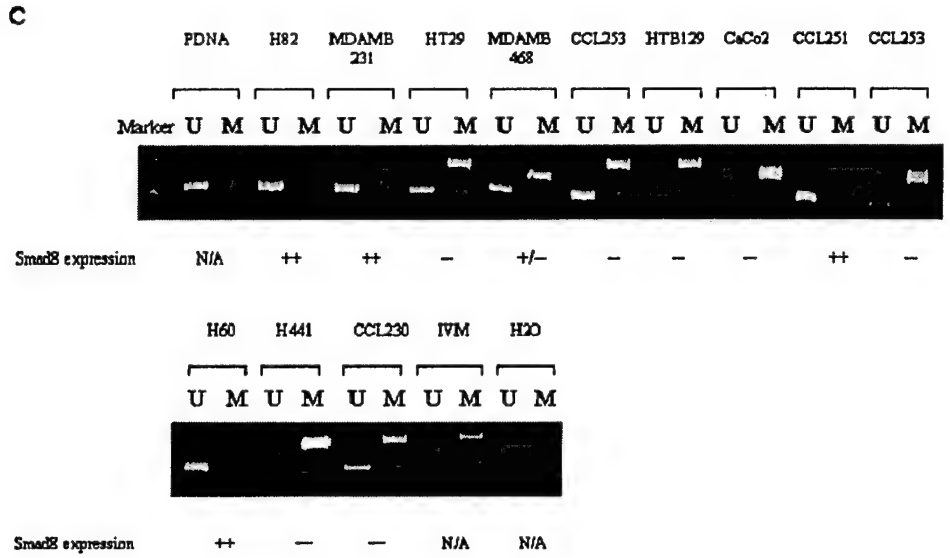

Figure 3C. Epigenetic gene silencing of the *Smad8* due to altered DNA methylation.

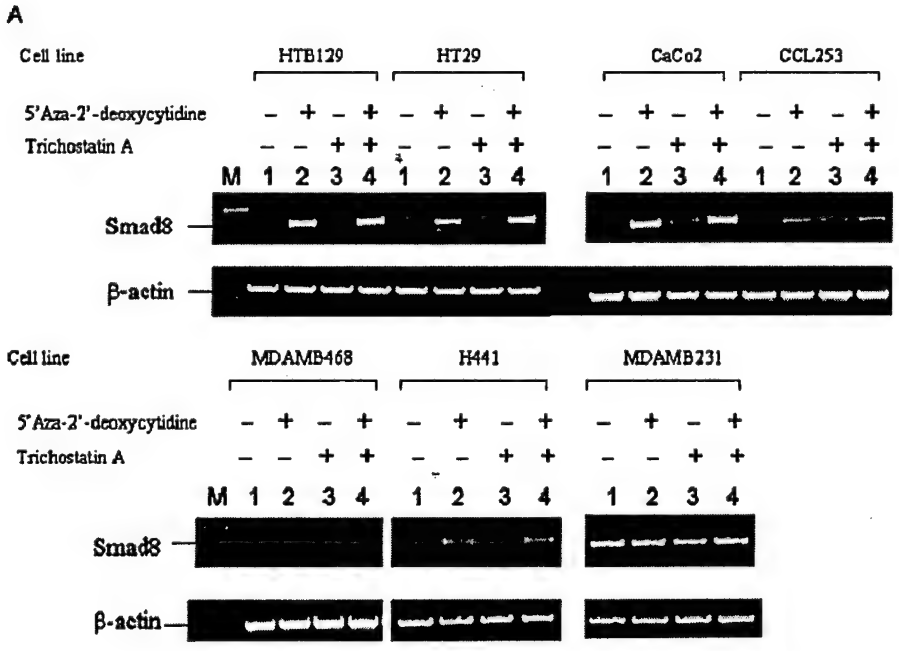

Figure 4A. The effects of DNA demethylation and inhibition of histone deacetylases on *Smad8* gene expression.

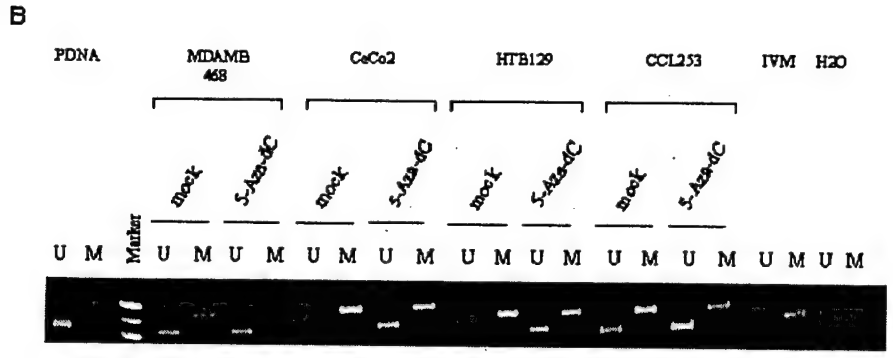

Figure 4B. The effects of DNA demethylation and inhibition of histone deacetylases on *Smad8* gene expression.

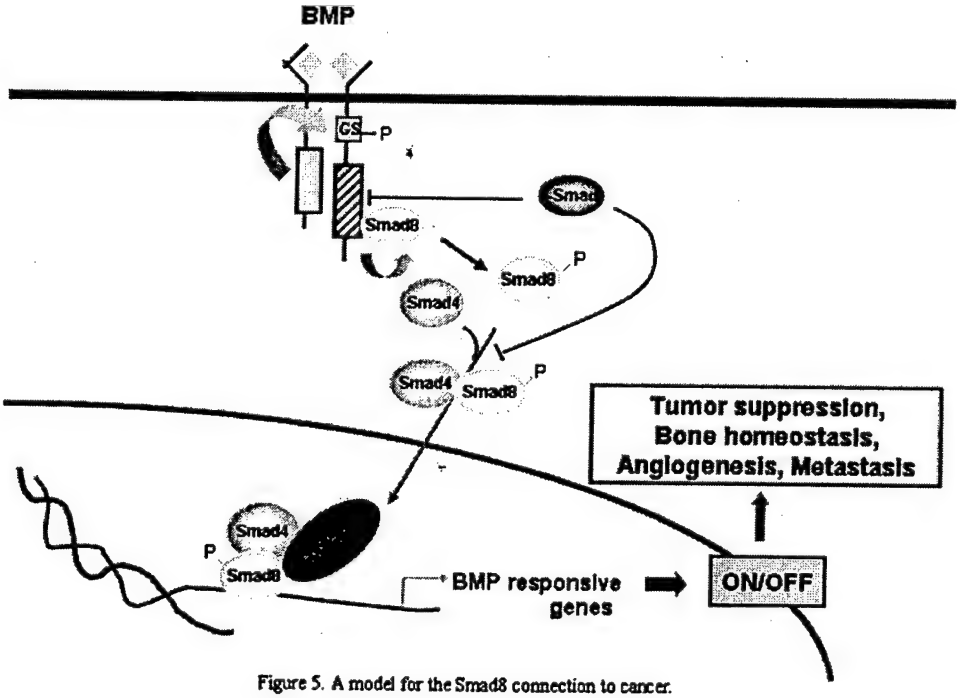

Figure 5. A model for the Smad8 connection to cancer.

Reprinted from **Epigenetics in Cancer Prevention**
Early Detection and Risk Assessment
Volume 983 of the *Annals of the New York Academy of Sciences*
March 2003

Histone Deacetylases: Unique Players in Shaping the Epigenetic Histone Code

SAM THIAGALINGAM,[a,b,c] KUANG-HUNG CHENG,[c] HYUNJOO J. LEE,[a] NORA MINEVA,[c] ARUNTHATHI THIAGALINGAM,[d] AND JOSE F. PONTE[a]

[a]*Genetics and Molecular Medicine Programs and Pulmonary Center, Department of Medicine,* [b]*Department of Genetics and Genomics, and* [c]*Department of Pathology and Laboratory Medicine, Boston University School of Medicine, Boston, Massachusetts 02118, USA*

[d]*Bayer Corporation, 333 Coney Street, East Walpole, Massachusetts 02032, USA*

ABSTRACT: **The epigenome is defined by DNA methylation patterns and the associated posttranslational modifications of histones. This histone code determines the expression status of individual genes dependent upon their localization on the chromatin. The silencing of gene expression is associated with deacetylated histones, which are often found to be associated with regions of DNA methylation as well as methylation at the lysine 4 residue of histone 3. In contrast, the activation of gene expression is associated with acetylated histones and methylation at the lysine 9 residue of histone 3. The histone deactylases play a major role in keeping the balance between the acetylated and deacetylated states of chromatin. Histone deacetylases (HDACs) are divided into three classes: class I HDACs (HDACs 1, 2, 3, and 8) are similar to the yeast RPD3 protein and localize to the nucleus; class II HDACs (HDACs 4, 5, 6, 7, 9, and 10) are homologous to the yeast HDA1 protein and are found in both the nucleus and cytoplasm; and class III HDACs form a structurally distinct class of NAD-dependent enzymes that are similar to the yeast SIR2 proteins. Since inappropriate silencing of critical genes can result in one or both hits of tumor suppressor gene (TSG) inactivation in cancer, theoretically the reactivation of affected TSGs could have an enormous therapeutic value in preventing and treating cancer. Indeed, several HDAC inhibitors are currently being developed and tested for their potency in cancer chemotherapy. Importantly, these agents are also potentially applicable to chemoprevention if their toxicity can be minimized. Despite the toxic side effects and lack of specificity of some of the inhibitors, progress is being made. With the elucidation of the structures, functions and modes of action of HDACs, finding agents that may be targeted to specific HDACs and potentially reactivate expression of only a defined set of affected genes in cancer will be more attainable.**

KEYWORDS: **histone deacetylases (HDAC); histone code; active histone code (AHC); silenced histone code (SHC); histone deacetylase inhibitor (HDACi); cancer therapy**

Address for correspondence: Sam Thiagalingam, Genetics Program, Boston University School of Medicine, 715 Albany Street, L320, Boston, MA 02118. Voice: 617-638-6013; fax: 617-638-4275.
samthia@bu.edu

Ann. N.Y. Acad. Sci. 983: 84–100 (2003).

INTRODUCTION

The impending completion of the Human Genome Project has led the scientific community to the cusp of identifying every gene within the DNA of our genome. However, many challenges still lie ahead, foremost of which may be deciphering the regulatory cues and mechanisms that allow these genes to be "turned on" or "off" depending upon the intra- and extracellular signals a cell receives. Eukaryotes have evolved a complex packaging of DNA that encumbers transcription, which requires accessible DNA to allow transcription factors and RNA polymerase to bind to promoters. Genomic DNA is packaged into highly ordered structures known as *chromatin*, which is composed of structural subunits called *nucleosomes*. Nucleosomes consist of 146 base pairs of DNA, which is the equivalent of two superhelical turns of DNA and an octamer of core histone proteins. The histones have numerous sites where posttranslational modifications can occur, and it has been proposed that the pattern of modifications acts as an information code that regulates processes that influence gene transcription. This pattern of modifications has been termed the *histone code*.[1,2] The particular pattern of histone modifications may play a role in determining the affinity for chromatin-associated proteins, which determine whether the chromatin takes on an active or silent state. DNA methylation and histone modification are the major contributors to chromatin modification, which, combined with ATP-dependent chromatin remodeling, is the principle epigenetic mechanism by which tissue-specific gene expression patterns and global gene silencing are established and maintained.[3]

This review will discuss one of the important chromatin-modifying effects, histone deacetylation, a process that is correlated with repression of gene expression. It is common knowledge that orderly expression of appropriate genes at optimal levels is central to the maintenance of the destined differentiated status of all the cells that make up the human body, while any alterations or inappropriate levels of gene expression may lead to cancer.[4] Therefore, we will also discuss the implications of the use of agents that influence gene expression by affecting histone acetylation as promising chemopreventive or chemotherapeutic agents in cancer.

THE HISTONE CODE

The basic building block of chromatin is the nucleosome, which consists of 146 base pairs of DNA wrapped around an octamer of histones represented by two copies each of histone (H) 2A, H2B, H3, and H4.[5,6] Histones are basic proteins that consist of a globular domain and an N-terminal tail that protrudes from the nucleosome. Although the histones are some of the most evolutionarily conserved proteins, they are also among the most variable in terms of posttranslational modifications. The histone tails emanating from the nucleosome are unstructured and serve as targets for characteristic covalent posttranslational modifications, including acetylation, phosphorylation, methylation, sumoylation, and ubiquitination. These posttranslational modifications determine the structure and pattern of chromatin condensation and determine the histone code involved in gene regulation.[7] Cytogenetic analysis of chromatin identified euchromatin and heterochromatin, where the heterochromatin is the portion that remains deeply stained (heteropyknotic) and highly condensed during

cell division. The heterochromatin region is generally rich in repetitive DNA sequences and very low in gene density. However, the extent of heterochromatin of specific regions may differ in different individuals or tissue types and may be determined by a complex process involving factors responsible for chromatin remodeling. The outcome of the chromatin remodeling process, the histone code, apparently determines a mechanistic basis not only for the spreading of heterochromatin but also for the epigenetic inheritance of the silent states of specific regions of chromatin. The location of a specific gene on the chromatin may eventually determine whether the gene is either expressed or silenced. This impact of gene location is known as position effect variegation (PEV). Heterochromatization of a formerly euchromatic region at its boundaries may have an enormous impact on the status of gene expression.

HUMAN HISTONE DEACETYLASES

In the mid-1960s, Allfrey and his colleagues were the first to observe histone acetylation and postulated that acetylation of core histones could regulate transcription.[8] Histone hyperacetylation correlated with increased transcription and hypoacetylation with repression. However, it was not until the early 1990s that the role of HDACs in this regulation came to prominence. The initial observations that implicated a role for HDACs in transcriptional regulation came from a screen to identify small molecules that could return spindle-like transformed NIH3T3 cells to the normal fibroblast-like morphology. An epoxyketone-containing cyclic tetrapeptide, trapoxin, was identified without the knowledge of what proteins this molecule was acting on.[9] Later, it was discovered that cells treated with trapoxin had hyperacetylated histones and that this molecule inhibited histone deacetylation.[10] It was not until 1996, however, that the protein target for trapoxin was identified with the cloning of the first histone deacetylase.[11] To date, 18 HDACs have been identified in humans, and their activities have been implicated in transcription,[12–17] cell cycle progression,[18–21] gene silencing,[22] differentiation,[23,24] DNA replication,[19,25–28] and the DNA damage response.[29–32] One question that often arises is why do humans need so many HDACs and what are the roles for each of these HDACs? Clues are starting to emerge, as well as many more questions. Grunstein and colleagues recently used microarray deacetylation maps in yeast to determine the genome-wide functions of yeast deacetylases. They showed that Rpd3 and Hda1 act predominantly on distinct promoters and gene classes and are recruited by novel mechanisms. Hda1 also deacetylates subtelomeric domains, which contain genes involved in gluconeogenesis, growth on nonglucose carbon sources, and adverse growth conditions. Sir2 was shown to deacetylate subtelomeric heterochromatin, while Hos1/Hos3 and Hos2 regulate ribosomal DNA and ribosomal protein genes.[33] Researchers have set out to delineate human HDAC-mediated events, and the one clear observation is that HDAC function is very complex. The number of HDACs, splice variants of these HDACs, proteins that associate with HDACs either alone or in multiprotein complexes, and posttranslational modifications such as phosphorylation and sumoylation all play a role in regulating the specificity of HDAC activity.

The HDACs can be separated into three classes based on their homology to yeast histone deacetylases (TABLE 1). Class I HDACs have high homology to the yeast

TABLE 1. Human histone deacetylase

Histone deacetylase	Amino acids	Sensitivity to TSA[a]	Chromosomal location	Reference
Class I				
HDAC1	482	yes	1p34.1	11, 131, 132
HDAC2	488	yes	6q21	131, 133
HDAC3	428	yes	5q31.1–5q31.3	34–36, 131
HDAC8	377	yes	Xq21.2–Xq21.3 or Xq13	37–39
Class II				
HDAC4	1084	yes	2q37	59, 68, 134
HDAC5	1122	yes	17q21	59, 60, 134
HDAC6	1215	yes	Xp11.23	59, 68, 134
HDAC7	952	yes	12q13.1	60
HDAC9	1011	yes	7p15–p21	60, 61, 68, 134
HDAC10	669	yes	22q13.31–13.33	62, 135–137
HDAC11[b]	347	yes	3p25.1	70
Class III (sirtuins)				
SIRT1	747	no	10q22.2	73
SIRT2	389	no	19q13	73, 76
SIRT3	399	no	11p15.5	73
SIRT4	314	no	12q	73
SIRT5	310	no	6p22.3	73
SIRT6	355	no	19p13.3	138
SIRT7	400	no	17q	138

[a]TSA, trichostatin A.
[b]HDAC11 has properties of both class I and class II HDACs.

RPD3 gene, whereas class II HDACs are homologous to the Hda1 gene. A third family, class III HDACs, were identified based on their similarity to the Sir2 gene.

Class I HDACS

The class 1 HDACs, HDAC1, HDAC2, HDAC3, and HDAC8,[34–39] all share a certain degree of homology to the yeast RPD3 gene, are around 400–500 amino acids long, generally localize to the nucleus, and are ubiquitously expressed in many human cell lines and tissues. All four members have a deacetylase catalytic domain, and HDAC1 and HDAC2 have a C terminal RB binding motif adjacent to a basic region. Each class I HDAC has been mapped to a chromosomal location (TABLE1); interestingly, HDAC 8 was shown to localize to Xq13 by FISH using the HDAC8 cDNA as the probe,[39] whereas another group using radiation hybrid mapping reported the location at Xq21.2–Xq21.3,[38] raising the possibility that a gene duplication event may have occurred. All four members have been shown to be sensitive to his-

tone deacetylase-specific inhibitors. Interestingly, the messenger RNAs of all but HDAC8 are upregulated in response to trichostatin A treatment, suggesting that HDAC inhibitors (HDACi) may trigger an autoregulatory loop that results in a compensatory feedback pathway.[40]

It is now becoming clear that these HDACs are parts of large protein complexes *in vivo* that direct gene-specific regulation of transcription, hormone signaling, the cell cycle, differentiation, and DNA repair. Class I HDACs have been shown to associate with the silencing mediator for the retinoid and thyroid hormone receptor complex (SMRT),[41] the CoREST complex, as well as the Sin3 and Mi-2/NuRD corepressor complexes.[42,43] HDAC1 and 2 are part of the core complex along with RbAp46/48. The Sin3 complex consists of this core complex in addition to SAP18 and 30, which aid in stabilizing the protein interactions; and mSin3A, which serves as the scaffold for the assembly of the complex.[44] The NuRD complex contains the core complex along with MTA2, CHD3, and CHD4, all of which contain DNA helicase/ATPase domains.[45] HDAC1 and 2 are found in the CoREST complex, but unlike the other complexes, neither RbAp46 nor RbAp48 is present. The remaining components are proteins homologous to MTA1 and 2, called CoREST and p110, respectively.[46] Members of the class I HDACs have also been found in association with Rb,[27,47] DNA methyltransferase 1,[48,49] TGIF/Smads,[50] glucocorticoid receptor,[51] and Sp1.[52] Recently, HDAC3 was shown to form a complex with N-CoR (nuclear receptor corepressor),[53,54] and this corepressor complex inhibits JNK activation through an integral subunit, GPS2.[55]

Recent work has implicated posttranslational modifications of HDAC in regulating HDAC activity and association potential. Galasinski *et al.* have shown that phosphorylated HDAC1 and 2 had a small increase in activity relative to that observed in the nonphosphorylated HDACs and that this increase was reversed upon phosphatase treatment.[56] These investigators went on to show that phosphorylation disrupted HDAC1 and 2 complex formation as well as the interaction between HDAC1 and mSin3 and YY1 but not RbAp46/48. Though HDAC1 has been shown to be phosphorylated by CK2, cAMP-dependent protein kinase, and protein kinase G *in vitro*, HDAC2 is uniquely phosphorylated by CK2. This HDAC2 phosphorylation promotes enzymatic activity and regulates complex formation, but has no effect on transcriptional repression.[57] David, Neptune, and DePinho have proposed another mechanism of regulation. They demonstrated that HDAC1 is a substrate for SUMO-1 (small ubiquitin-related modifier) modification and that mutations in the target residues reduced transcriptional repression without affecting the ability of HDAC1 to associate with mSin3.[58] These observations suggest that SUMO-1 modification regulates the biological effects of HDAC1 by potentiating its histone deacetylase activity.

Class II HDACs

Once the novel yeast deacetylase Hda1 was characterized, several groups simultaneously isolated some of the human homologues using database searches. From this, HDAC4, 5, 6, and 7 were identified.[59,60] Subsequently, HDAC9[61] and HDAC10[62] were isolated and assigned to class II. These HDACs are twice as large (~1000 amino acids) as the class I family members, and most have a COOH terminus catalytic domain, except for HDAC6, which has a second catalytic domain in the

NH_2 terminus. HDAC10 has an NH_2 terminus catalytic domain and a COOH terminus pseudorepeat that shares homology with the catalytic domain. Class II HDACs are also sensitive to HDACi; but, unlike class I HDACs, class II HDACs are cytoplasmic and are shuttled to the nucleus as they are needed. HDAC10 is an exception, as it has been shown to be a nuclear protein. Class II HDACs are also differentially expressed in human tissue, with the highest levels being found in the heart, brain, and skeletal muscle.[24,43]

Class II HDACs have also been shown to be a part of larger multiprotein complexes. HDAC4 and 5 associate with HDAC3[59] and form a complex with N-CoR and SMRT.[63] The association with HDAC3 has been shown to be regulated by 14-3-3. Interaction of HDAC4 or 5 with 14-3-3 proteins sequesters the protein in the cytoplasm. When this interaction is lost, HDAC4 and 5 translocate to the nucleus and associate with HDAC3 and repress gene expression.[64] A similar mechanism has been proposed for the regulation of the importin-α-HDAC4 association by 14-3-3.[65] A recent study demonstrated that the catalytic domain of HDAC4 interacts with HDAC3 through N-CoR/SMRT. The authors of this study suggest that class II HDACs regulate transcription by bridging the SMRT/N-CoR-HDAC3 complex and select transcription factors independently of HDAC activity.[66] The recently identified HDAC10 also interacts with SMRT as well as with HDAC2.[62]

A common NH_2 terminal extension in HDAC4, 5, and 7 allows them to interact with the MEF2 family of transcription factors once they translocate from the cytoplasm to the nucleus. These interactions play an important role in activating muscle-specific genes and differentiation in both smooth and skeletal muscle.[67,68] Class II HDACs have also been reported to interact with the COOH terminal binding protein (CtBP) and repress MEF2-mediated transcription.[69]

The 11th member of the HDAC family was recently cloned and characterized; interestingly, it has properties seen in both classes of HDACs.[70] The protein is 347 amino acids long, with homology in the core catalytic domains to both class I and class II HDACs. The size of the protein is in line with class I HDACs, but HDAC11 is differentially expressed in the heart, brain, skeletal muscle, and kidney, which is typical of class II HDACs. The protein is predominantly nuclear, and like its family members, HDAC11 is sensitive to HDACi. HDAC11 associated with complexes that contained HDAC6,[70] which has recently been shown to function as a tubulin deacetylase.[71]

Class III HDACs

The third family of histone deacetylases, sirtuins, are homologues of the yeast Sir2 gene, which has been implicated in chromatin silencing, cellular metabolism, and aging.[72] There are seven sirtuins in humans, SIRT1–7, most of which average around 300–400 amino acids, except for SIRT1 which has 747 (TABLE 1). The catalytic domains average 275 amino acids and contain two CXXC motifs that function as zinc finger domains[73] and at least one hydrophobic region that potentially functions as a leucine zipper.[74] The histone deacetylase activity of these enzymes is dependent upon NAD^+,[75] and the yeast Sir2 has intrinsic ADP-ribosyltransferase activity.[73] Mutational analysis indicated that Gly 270 and Asn 345 are critical amino acids whereby deacetylase activity was abolished in the 345 mutant and diminished in the 270 mutant.[75]. ADP-ribosyltransferase activity was also abolished in the 345

mutant and severely decreased in the 270 mutant.[75] Immunofluorescence studies have demonstrated that unlike the yeast Sir2, human Sir2 does not localize in the nucleus.[76] Though the field of human sirtuins is in its infancy, some interesting developments are starting to emerge, the foremost being the association of SIRT1 with p53. SIRT1 has been shown to specifically associate with and deacetylate p53, thereby repressing p53-mediated transcriptional activation, which prevents growth inhibition or apoptosis in response to DNA damage.[30,31,77–79] These findings could have a tremendous impact on p53-based cancer therapy, as inhibitors of SIRT1 could be used in combination with current therapeutic protocols to enhance efficacy.

THE EPIGENOME AND ACETYLATION

Epigenetic changes of the genome include DNA methylation and modifications of histones. In humans, DNA cytosine methyltransferases (Dnmt1, Dnmt3a, Dnmt3b) usually add a methyl group to the 5′-carbon of a cytosine located next to a guanine (5′-CpG-3′). These CpG sequences are found in islands mainly in the 5′-regions such as the promoter, first exon, and sometimes in the first intron of housekeeping genes as well as tissue-specific genes. Although most CpG islands are unmethylated in normal cells, they could become methylated during development, differentiation, or cancer and play a part in gene regulation. Among the histone modifications, acetylation of core histone tails has been shown to be dependent on the opposing activities of two types of enzymes, histone acetyltransferases (HATs) and histone deacetylases (HDACs). HATs acetylate the ε-groups of the lysine residues of the histone tails, and their removal by HDACs restores the positive charge on these residues. Actively transcribed regions of the chromatin are generally enriched with highly acetylated histones H3 and H4 in euchromatic regions of the genome. Methylation of histones by proteins bearing the SET (Su[var], Enhancer of zestes, trithorax) domain also targets lysine residues. Distinct methyl transferases (H3-K4 methyltransferase and H3-K9 methyltransferases such as Suv39h1, Suv39h2, G9a, ESET/SetDB1 and Eu-HMTase) methylate histone H3 either at lysine 4 (H3-meK4), lysine 9 (H3-meK9), or other lysine residues. The regions of chromatin with H3-meK4 modifications usually harbor lysine 9 modified by acetylation (H3-AcK9), marking active euchromatin; while the presence of H3-meK9 is correlated with condensed heterochromatin.[80,81] The chromodomain of HP1 (heterochromatin protein 1) binds to H3-meK9 with high affinity and is involved in heterochromatin assembly through the oligomerization of the HP1 proteins.[82,83] Furthermore, H3-meK4 inhibits the binding of the nucleosome remodeling deacetylase (NuRD) repressor complex to H3 histone tails to ensure disruption of the silencing process by protein–protein interactions, thereby resulting in expression of otherwise silent genes.

The multisubunit complex with ATPase activity known as SWI/SNF in human cells consists of either of the two ATPase subunits, BRG1 (a human homologue of the yeast Swi2/Snf2) or hBRM. The human SWI/SNF complexes play a major role in chromatin remodeling and are not only enriched in active chromatin but also present and found to form complexes with corepressors such as Sin3, HDAC1, HDAC2, HDAC3, N-CoR, and KAP1 (krab-associated protein 1). These observations suggest that SWI/SNF plays important roles in both regulation of transcription and gene repression.[84] Prominent examples illustrating the differential effects of

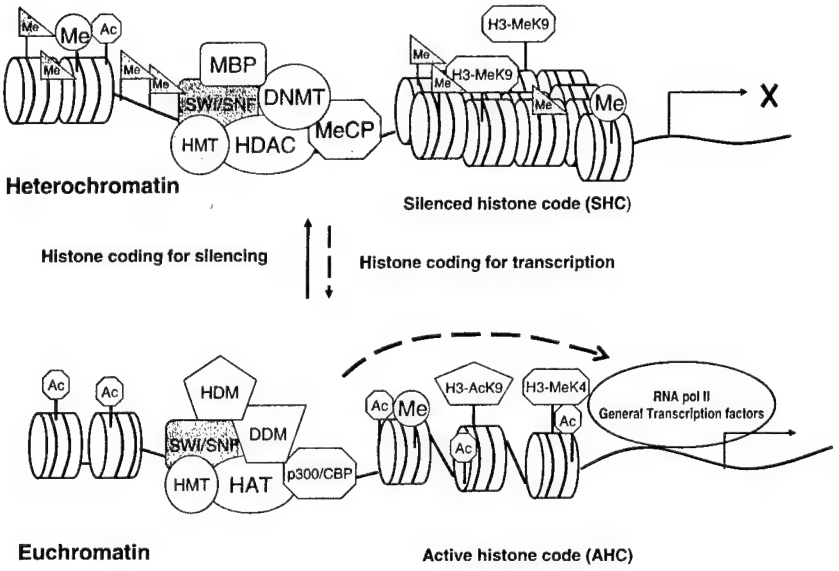

FIGURE 1. The role of HDACs in the histone code. The molecular details of the various modifications of the epigenome in relation to the heterochromin and euchromatin can be found in the text. Abbreviations: HAT, histone acetyltransferase; HDAC, histone deacetylase; MBP, Methyl Binding Protein ; MeCP, methyl-CpG binding protein; H3-meK4, histone H3 methylated at lysine 4; H3-meK9, histone H3 methylated at lysine 9; H3-AcK9, histone H3 acetylated at lysine 9; DNMT, DNA methyl transferase; SWI/SNF, chromatin remodeling multiprotein complex with ATPase activity; Me in a *triangle* denotes DNA CpG methylation; Me in a *circle* denotes histone methylation; Ac , histone acetylation; HMT, histone methyltransferase; p300/CBP, CREB binding protein; DDM, DNA demethylase; HDM, histone demethylase.

SWI/SNF are the interaction of the retinoblastoma (Rb) tumor suppressor protein (1) with both BRG1 and hBRM to form a hSWI/SNF repressor complex regulating expression of cyclins and cyclin-dependent kinases (cdks) during S phase and (2) with HDACs to repress certain genes such as cyclin E during the G1 phase.[85–89] Interestingly, studies with the filamentous fungi *Neurospora crassa* showed that mutation of H3-lys9 resulted in a loss of DNA methylation *in vivo,* suggesting that H3-lys9 methylation could be coupled to DNA methylation in other organisms and that a similar mechanism may play a role in silencing chromatin in mammals.

These studies lead us to believe that DNA methyltransferases might be taking cues from the histone code. The two repression mechanisms, DNA methylation and histone deacetylation, are apparently connected by the methyl-CpG binding proteins (MBPs), such as MeCP2, MBD1, MBD2, MBD3, MBD4, and Kaiso or DNA methyl transferases. The MeCP2 protein can interact with histone deacetylases (HDAC1 and HDAC2) via the corepressor Sin3. On the other hand, MBD2 is also associated with HDAC1, which interacts with the Sin3 or NuRD complex.[44] Other MBPs are also believed to recruit HDAC activity.[44,90] Direct interactions, as well as interac-

tions in a complex between the various DNA methyltransferases—Dnmt1, Dnmt3a, and Dnmt3b as well as a Dnmt3 family homologue, DnmtL—have been demonstrated.[49,91,92] These observations strongly suggest that the distribution of the acetylation of histones in the chromatin resulting from the complex nature of the epigenome marks the status of gene expression (FIG. 1).

HDAC INHIBITORS AS CANCER PREVENTIVE AND THERAPEUTIC AGENTS

HDAC inhibitors cause the accumulation of acetylated histones in nucleosomes, which results in the expression of a specific set of genes that can lead to cell arrest, differentiation, or apoptosis. Therefore, they have the potential for use in the chemoprevention and treatment of cancer.[49,91–94] Inhibitors of HDACs have been isolated from natural sources as well as derived from synthetic compounds, as summarized in TABLE 2. Many different structural classes of HDAC inhibitors have been reported, including:

(1) short-chain fatty acids—e.g., sodium n-butyrate (NaBu);[95,96]
(2) hydroxamic acids, such as trichostatin A (TSA,)[97] suberoylanilide hydroxamic acid (SAHA,)[98–100] and Oxamflatin;[101]
(3) cyclic tetrapeptides containing a 2-amino-8-oxo-9,10-epoxy-decanoyl (AOE) moiety—e.g., trapoxin;[10,102]
(4) cyclic tetrapeptides without an AOE moiety—e.g. apicidin[103,104] and FR901228[105]; and
(5) benzamides—e.g., MS-27-275.[106,107]

Sodium n-butyrate (NaBu) is a nonspecific inhibitor, which has been shown to reduce the proliferation of many tumor cell lines, enhance diffrentiation, and stimulate apoptosis, leading to decreased viability of cells.[95,96,108, 109] The butyrates are the only class, to date, that have been approved for clinical use; but they are far from ideal inhibitors, as they are nonspecific, exerting effects on multiple enzyme systems, and the dose required to inhibit deacetylation is in the millimolar range. A number of investigators have shown that sodium butyrate enhances the efficacy of retinoic acid (RA) in a number of cell lines, including the S91 melanoma line.[110] TSA, originally developed as an antifungal agent, is a potent and reversible inhibitor of histone deacetylase; nanomolar concentrations of it inhibit deacetylase activity, targeting the cell cycle progression of several cell types, inducing cell growth arrest at both the G1 and G2/M phases, and in some cases also inducing apoptosis.[18,96,97,111–113] TSA inhibition of HDACs has been shown to alter gene expression (twofold increase or decrease) in roughly 2% of expressed genes, suggesting that the action of TSA is selective.[114] Similar results were also observed in transformed cultured cells treated with SAHA.[93] SAHA[100] is a cell-permeable inhibitor of HDACs that structurally resembles TSA. SAHA has been shown to induce growth inhibition,[99,115] differentiation,[100,116] and apoptosis in a variety of cell types, including ARP-1 multiple myeloma cells, the LNCaP prostate cancer cell line, and U937 leukemia cells.[93] SAHA also induces caspase-dependent apoptosis and downregulation of daxx in acute promyelocytic leukemia with t(15;17),[117] as well as antiangiogenesis activity by altering VEGF signaling in HUVEC cells.[118] In *in vivo* studies, the incidence of

TABLE 2. HDAC inhibitors

Class of inhibitor	Example	Optimal concentration	Reference
Short chain fatty acid	butyrates	1.5 mM	93, 139
Hydroxamic acids	trichostatin A	40–70 nM	102
	SAHA	2–5 μM	100
	oxamflatin	μM range	101
Cyclic tetrapeptides with AOE moiety	trapoxin A	50 nM	10
Cyclic peptides without AOE moiety	FR901228	μM range	105
	apicidin	2–4 nM	104
Benzamides	MS-27-275	2–5 μM	107

mammary tumors was reduced by 40% and the mean tumor volume by 78% without any side effects when rats with methylnitrosourea-induced mammary carcinomas were fed SAHA (900 parts/million.)[99] Two other studies also revealed inhibition of tumor growth by SAHA in mice with lung cancer induced by administration of 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone and in nude mice transplanted with CWR22 androgen-dependent prostate cancer.[93]

Other HDACi shown to inhibit tumor growth in animal models include oxamflatin, MS-27-275, and azeloic bishydroxamate. In many of these cases, no toxicity, evaluated by weight gain and histologic examination, was observed. Trapoxin (TPX) [cyclo-(L-phenylalanyl-L-phenylalanyl-D-pipecolinyl-L-2-amino-8-oxo-9,10-epoxy-decanoyl)] is a fungal product that can induce morphological reversion of transformed NIH3T3 fibroblasts. Removing an epoxide group in trapoxin completely abolished the inhibitory activity, which suggests that trapoxin binds covalently to the histone deacetylase via the epoxide group. Trichostatin A reversibly inhibits HDACs, whereas trapoxin causes inhibition by irreversible binding to the HDAC. However, they have been shown to induce nearly identical biological effects on the cell cycle and differentiation.[102,119] Apicidin [cyclo(*N*-*O*-methyl-L-tryptophanyl-L-isoleucinyl-D-pipecolinyl-L-2-amino-8-oxodecanoyl)] is a fungal metabolite shown to inhibit both mammalian and protozoan HDACs ($IC_{50} = 0.2–1.5$ nM). Apicidin can lead to a morphological reversal and growth inhibition of H-ras MCF10A cells similar to that induced by other HDAC inhibitors.[103] The growth inhibition of apicidin on HeLa cells is accompanied by morphological changes, cell cycle arrest at the G1 phase with increased induction of p21/WAF1/Cip1 and decreased phosphorylation of the Rb protein, and accumulation of hyperacetylated histone H4.[103,104] In another study, apicidin was shown to induce apoptosis and Fas/Fas ligand expression in the human acute promyelocytic leukemia cells HL60.[120] The newly synthesized benzamide derivative of MS-27-275 can induce p21 (WAF1/CIP1) and gelsolin, resulting in an altered cell cycle distribution.[106,107] In some studies, an increase in the accumulation of acetylated histones H3 and H4 was detected in the TβR II promoter after treatment with MS-275, and MS-27-275 was able to induce an increase in TGF-betaRII mRNA to restore TGF-beta signaling.[121] HDACi have also been used in combinational therapy most frequently with retinoic acids in hematological

cancers.[110,122–127] This area of study is more thoroughly described in a review by Pandolfi.[128]

The mechanisms of inhibition of HDACs by these inhibitors are coming to light with the resolution of the structure of the catalytic core of the HDACs.[129] HDACs have a homologous 390–amino acid catalytic core, and the residues that form the active site are conserved across all HDACs. An HDAC homologue in *Aquifex aeolicus* called HDLP was used in crystallography studies to analyze an HDLP-TSA and HDLP-SAHA complex. These studies revealed that a tubular pocket, a zinc binding site, and two asparagine-histidine charge relay systems form the active catalytic site of HDLP. The hydroxamic acid moieties of TSA and SAHA bind to the zinc in the tubular pocket, and this interaction is believed to be critical in inhibiting the enzyme.[93,129]

Recently, there have been reports on inhibitors of sirtuins, which in general are not inhibited by these types of HDACi. Identifying and generating inhibitors to this class of HDACs would expedite the dissection of their biological functions and, in the long term, could possibly be used in combinational therapy, especially in light of the interaction between SIRT1 and p53. Nonhydrolyzable NAD analogues have been used, but they are problematic in that they nonspecifically inhibit other NAD-dependent enzymes. Small molecules that contain a 2-hydroxyl-1-napthol moiety have been developed and have been shown to inhibit sirtuins.[130] These compounds may be the building blocks of an approach to find specific inhibitors to each of the sirtuins, allowing the delineation of the role of these proteins in transcriptional regulation, cell growth, DNA repair, apoptosis, and development.

CONCLUSIONS AND FUTURE PERSPECTIVES

It is becoming evident that the key to effectively using the information provided by the Human Genome Project hinges on the accurate interpretation of the histone code. The roles for HDACs in the histone code and transcriptional regulation are becoming clearer, but the identification of splice variants of some of the HDACs and posttranslational modifications to the HDACs shows just how complex the regulation of these enzymes and the complexes that they are found in can be. The ever growing list of HDACi will help to elucidate the roles of these HDACs in mediating growth arrest, differentiation, and cell death. The identification of HDACi specific to HDACs involved in these processes through regulation of expression of a defined set of genes affected in cancer would be of great value in cancer prevention and therapy and will continue to be a major focus of research in these fields.

ACKNOWLEDGMENTS

Work in the authors' laboratory is supported by grants from the NIH (RO1 ES10377) and the Evans Medical Foundation, and a Career Development Award from the Department of Defense (DAMD 17-01-1-0160). S.T. is a Dolphin Trust Investigator of the Medical Foundation. H.J.L. is a Karin Grunebaum Fellow and is supported by NIH Training Grant (T32HL07035).

REFERENCES

1. STRAHL, B.D. & C.D. ALLIS. 2000. The language of covalent histone modifications. Nature **403:** 41–45.
2. RICE, J.C. & C.D. ALLIS. 2001. Code of silence. Nature **414:** 258–261.
3. Li, E. 2002. Chromatin modification and epigenetic reprogramming in mammalian development. Nat. Rev. Genet. **3:** 662–673.
4. THIAGALINGAM, S. *et al.* 2002. Loss of heterozygosity as a predictor to map tumor suppressor genes in cancer: molecular basis of its occurrence. Curr. Opin. Oncol. **14:** 65–72.
5. LUGER, K. & T.J. RICHMOND. 1998. The histone tails of the nucleosome. Curr. Opin. Genet. Dev. **8:** 140–146.
6. LUGER, K. *et al.* 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature **389:** 251–260.
7. JENUWEIN, T. & C.D. ALLIS. 2001. Translating the histone code. Science **293:** 1074–1080.
8. ALLFREY, V.G. 1966. Structural modifications of histones and their possible role in the regulation of ribonucleic acid synthesis. Proc. Can. Cancer Conf. **6:** 313–335.
9. ITAZAKI, H. *et al.* 1990. Isolation and structural elucidation of new cyclotetrapeptides, trapoxins A and B, having detransformation activities as antitumor agents. J. Antibiot. (Tokyo) **43:** 1524–1532.
10. KIJIMA, M. *et al.* 1993. Trapoxin, an antitumor cyclic tetrapeptide, is an irreversible inhibitor of mammalian histone deacetylase. J. Biol. Chem. **268:** 22429–22435.
11. TAUNTON, J., C.A. HASSIG & S.L. SCHREIBER. 1996. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science **272:** 408–411.
12. ZHANG, C.L. *et al.* 2002. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. Cell **110:** 479–488.
13. SUN, J.M. *et al.* 2002. The transcriptional repressor Sp3 is associated with CK2-phosphorylated histone deacetylase 2. J. Biol. Chem. **277:** 35783–35786.
14. WON, J., J. YIM & T.K. KIM. 2002. Sp1 and Sp3 recruit histone deacetylase to repress transcription of human telomerase reverse transcriptase (hTERT) promoter in normal human somatic cells. J. Biol. Chem. **277:** 38230–38238.
15. YU, Z., W. ZHANG & B.C. KONE. 2002. Histone deacetylases augment cytokine induction of the iNOS gene. J. Am. Soc. Nephrol. **13:** 2009–2017.
16. ZHANG, Y. & M.L. DUFAU. 2002. Silencing of transcription of the human luteinizing hormone receptor gene by histone deacetylase-mSin3A complex. J. Biol. Chem. **277:** 33431–33438.
17. OZAWA, Y. *et al.* 2001. Histone deacetylase 3 associates with and represses the transcription factor GATA-2. Blood **98:** 2116–2123.
18. KIM, Y.B. *et al.* 2000. Mechanism of cell cycle arrest caused by histone deacetylase inhibitors in human carcinoma cells. J. Antibiot. (Tokyo) **53:** 1191–1200.
19. MILUTINOVIC, S., Q. ZHUANG & M. SZYF. 2002. Proliferating cell nuclear antigen associates with histone deacetylase activity, integrating DNA replication and chromatin modification. J. Biol. Chem. **277:** 20974–20978.
20. Wang, C. *et al.* 2001. Histone acetylation and the cell-cycle in cancer. Front. Biosci. 6: D610–629.
21. ZHANG, Z.K. *et al.* 2002. Cell cycle arrest and repression of cyclin D1 transcription by INI1/hSNF5. Mol. Cell. Biol. **22:** 5975–5988.
22. BESTOR, T.H. 1998. Gene silencing. Methylation meets acetylation. Nature **393:** 311–312.
23. KIM, J. *et al.* 1999. Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. Immunity **10:** 345–355.
24. ZHOU, X. *et al.* 2000. Identification of a transcriptional repressor related to the noncatalytic domain of histone deacetylases 4 and 5. Proc. Natl. Acad. Sci. USA **97:** 1056–1061.
25. XIE, A.Y. & W.R. FOLK. 2002. Inhibition of polyomavirus ori-dependent DNA replication by mSin3B. J. Virol. **76:** 11809–11818.
26. GRUFFAT, H., E. MANET & A. SERGEANT. 2002. MEF2-mediated recruitment of class II HDAC at the EBV immediate early gene BZLF1 links latency and chromatin remodeling. EMBO Rep. **3:** 141–146.

27. LAI, A. *et al.* 2001. RBP1 recruits the mSIN3-histone deacetylase complex to the pocket of retinoblastoma tumor suppressor family proteins found in limited discrete regions of the nucleus at growth arrest. Mol. Cell. Biol. **21:** 2918–2932.
28. JOHNSON, C.A. *et al.* 2001. Deacetylase activity associates with topoisomerase II and is necessary for etoposide-induced apoptosis. J. Biol. Chem. **276:** 4539–4542.
29. JUAN, L.J. *et al.* 2000. Histone deacetylases specifically down-regulate p53-dependent gene activation. J. Biol. Chem. **275:** 20436–20443.
30. LUO, J. *et al.* 2000. Deacetylation of p53 modulates its effect on cell growth and apoptosis. Nature **408:** 377–381.
31. LUO, J. *et al.* 2001. Negative control of p53 by Sir2alpha promotes cell survival under stress. Cell **107:** 137–148.
32. CHEN, Z. *et al.* 2002. Induction and superinduction of growth arrest and DNA damage gene 45 (GADD45) alpha and beta messenger RNAs by histone deacetylase inhibitors trichostatin A (TSA) and butyrate in SW620 human colon carcinoma cells. Cancer Lett. **188:** 127–140.
33. ROBYR, D. *et al.* 2002. Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. Cell **109:** 437–446.
34. YANG, W.M. *et al.* 1997. Isolation and characterization of cDNAs corresponding to an additional member of the human histone deacetylase gene family. J. Biol. Chem. **272:** 28001–28007.
35. DANGOND, F. *et al.* 1998. Differential display cloning of a novel human histone deacetylase (HDAC3) cDNA from PHA-activated immune cells. Biochem. Biophys. Res. Commun. **242:** 648–652.
36. EMILIANI, S. *et al.* 1998. Characterization of a human RPD3 ortholog, HDAC3. Proc. Natl. Acad. Sci. USA **95:** 2795–2800.
37. HU, E. *et al.* 2000. Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor. J. Biol. Chem. **275:** 15254–15264.
38. BUGGY, J.J. *et al.* 2000. Cloning and characterization of a novel human histone deacetylase, HDAC8. Biochem. J. **350** Pt 1: 199–205.
39. VAN DEN WYNGAERT, I. *et al.* 2000. Cloning and characterization of human histone deacetylase 8. FEBS Lett. **478:** 77–83.
40. DANGOND, F. & S.R. GULLANS. 1998. Differential expression of human histone deacetylase mRNAs in response to immune cell apoptosis induction by trichostatin A and butyrate. Biochem. Biophys. Res. Commun. **247:** 833–837.
41. KAO, H.Y. *et al.* 1998. A histone deacetylase corepressor complex regulates the Notch signal transduction pathway. Genes Dev. **12:** 2269–2277.
42. VERDEL, A. *et al.* 2000. Active maintenance of mHDA2/mHDAC6 histone-deacetylase in the cytoplasm. Curr. Biol. **10:** 747–749.
43. CRESS, W.D. & E. SETO. 2000. Histone deacetylases, transcriptional control, and cancer. J. Cell. Physiol. **184:** 1–16.
44. ZHANG, Y. *et al.* 1999. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev. **13:** 1924–1935.
45. KNOEPFLER, P.S. & R.N. EISENMAN. 1999. Sin meets NuRD and other tails of repression. Cell **99:** 447–450.
46. YOU, A. *et al.* 2001. CoREST is an integral component of the CoREST- human histone deacetylase complex. Proc. Natl. Acad. Sci. USA **98:** 1454–1458.
47. LAI, A. *et al.* 1999. RBP1 recruits both histone deacetylase-dependent and -independent repression activities to retinoblastoma family proteins. Mol. Cell. Biol. **19:** 6632–6641.
48. ROBERTSON, K.D. *et al.* 2000. DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nat. Genet. **25:** 338–342.
49. ROUNTREE, M.R., K.E. BACHMAN & S.B. BAYLIN. 2000. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat. Genet. **25:** 269–277.
50. WOTTON, D. *et al.* 1999. A Smad transcriptional corepressor. Cell **97:** 29–39.
51. ITO, K., P.J. BARNES & I.M. ADCOCK. 2000. Glucocorticoid receptor recruitment of histone deacetylase 2 inhibits interleukin-1beta-induced histone H4 acetylation on lysines 8 and 12. Mol. Cell. Biol. **20:** 6891–6903.

52. CHOI, H.S. *et al.* 2002. Trichostatin A, a histone deacetylase inhibitor, activates the IGFBP-3 promoter by upregulating Sp1 activity in hepatoma cells: alteration of the Sp1/Sp3/HDAC1 multiprotein complex. Biochem. Biophys. Res. Commun. **296:** 1005–1012.
53. WEN, Y.D. *et al.* 2000. The histone deacetylase-3 complex contains nuclear receptor corepressors. Proc. Natl. Acad. Sci. USA **97:** 7202–7207.
54. LI, J. *et al.* 2000. Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. EMBO J. **19:** 4342–4350.
55. ZHANG, J. *et al.* 2002. The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. Mol. Cell. **9:** 611–623.
56. GALASINSKI, S.C. *et al.* 2002. Phosphatase inhibition leads to histone deacetylases 1 and 2 phosphorylation and disruption of corepressor interactions. J. Biol. Chem. **277:** 19618–19626.
57. TSAI, S.C. & E. SETO. 2002. Regulation of histone deacetylase 2 by protein kinase CK2. J. Biol. Chem. **277:** 31826–31833.
58. DAVID, G., M.A. NEPTUNE & R.A. DEPINHO. 2002. SUMO-1 modification of histone deacetylase 1 (HDAC1) modulates its biological activities. J. Biol. Chem. **277:** 23658–23663.
59. GROZINGER, C.M., C.A. HASSIG & S.L. SCHREIBER. 1999. Three proteins define a class of human histone deacetylases related to yeast Hda1p. Proc. Natl. Acad. Sci. USA **96:** 4868–4873.
60. FISCHLE, W. *et al.* 1999. A new family of human histone deacetylases related to Saccharomyces cerevisiae HDA1p. J. Biol. Chem. **274:** 11713–11720.
61. ZHOU, X. *et al.* 2001. Cloning and characterization of a histone deacetylase, HDAC9. Proc. Natl. Acad. Sci. USA **98:** 10572–10577.
62. FISCHER, D.D. *et al.* 2002. Isolation and characterization of a novel class II histone deacetylase, HDAC10. J. Biol. Chem. **277:** 6656–6666.
63. HUANG, E.Y. *et al.* 2000. Nuclear receptor corepressors partner with class II histone deacetylases in a Sin3-independent repression pathway. Genes Dev. **14:** 45–54.
64. GROZINGER, C.M. & S.L. SCHREIBER. 2000. Regulation of histone deacetylase 4 and 5 and transcriptional activity by 14-3-3-dependent cellular localization. Proc. Natl. Acad. Sci. USA **97:** 7835–7840.
65. BANNISTER, A.J. *et al.* 2000. Acetylation of importin-alpha nuclear import factors by CBP/p300. Curr. Biol. **10:** 467–470.
66. FISCHLE, W. *et al.* 2002. Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR. Mol Cell. **9:** 45–57.
67. LEMERCIER, C. *et al.* 2000. mHDA1/HDAC5 histone deacetylase interacts with and represses MEF2A transcriptional activity. J. Biol. Chem. **275:** 15594–15599.
68. MISKA, E.A. *et al.* 1999. HDAC4 deacetylase associates with and represses the MEF2 transcription factor. EMBO J. **18:** 5099–5107.
69. ZHANG, C.L. *et al.* 2001. Association of COOH-terminal-binding protein (CtBP) and MEF2-interacting transcription repressor (MITR) contributes to transcriptional repression of the MEF2 transcription factor. J. Biol. Chem. **276:** 35–39.
70. GAO, L. *et al.* 2002. Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family. J. Biol. Chem. **277:** 25748–25755.
71. HUBBERT, C. *et al.* 2002. HDAC6 is a microtubule-associated deacetylase. Nature **417:** 455–458.
72. GUARENTE, L. 2000. Sir2 links chromatin silencing, metabolism, and aging. Genes Dev. **14:** 1021–1026.
73. FRYE, R.A. 1999. Characterization of five human cDNAs with homology to the yeast SIR2 gene: Sir2-like proteins (sirtuins) metabolize NAD and may have protein ADP-ribosyltransferase activity. Biochem. Biophys. Res. Commun. **260:** 273–279.
74. BRACHMANN, C.B. *et al.* 1995. The SIR2 gene family, conserved from bacteria to humans, functions in silencing, cell cycle progression, and chromosome stability. Genes Dev. **9:** 2888–2902.
75. IMAI, S. *et al.* 2000. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. Nature **403:** 795–800.

76. AFSHAR, G. & J. P. MURNANE. 1999. Characterization of a human gene with sequence homology to Saccharomyces cerevisiae SIR2. Gene **234:** 161–168.
77. KUZMICHEV, A. *et al.* 2002. Role of the Sin3-histone deacetylase complex in growth regulation by the candidate tumor suppressor p33(ING1). Mol. Cell. Biol. **22:** 835–848.
78. LANGLEY, E. *et al.* 2002. Human SIR2 deacetylates p53 and antagonizes PML/p53-induced cellular senescence. EMBO J. **21:** 2383–2396.
79. VAZIRI, H. *et al.* 2001. hSIR2(SIRT1) functions as an NAD-dependent p53 deacetylase. Cell **107:** 149–159.
80. BANNISTER, A.J. *et al.* 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature **410:** 120–124.
81. SCHNEIDER, R., A.J. BANNISTER & T. KOUZARIDES. 2002. Unsafe SETs: histone lysine methyltransferases and cancer. Trends Biochem. Sci. **27:** 396–402.
82. BOEKE, J. *et al.* 2000. The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A. J. Biol. Chem. **275:** 34963–34967.
83. LACHNER, M. *et al.* 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature **410:** 116–120.
84. NEELY, K. & J. WORKMAN. 2002. The complexity of chromatin remodeling and its links to cancer. Biochim. Biophys. Acta **1603:** 19.
85. STROBECK, M.W. *et al.* 2000. BRG-1 is required for RB-mediated cell cycle arrest. Proc. Natl. Acad. Sci. USA **97:** 7748–7753.
86. TROUCHE, D. *et al.* 1997. RB and hbrm cooperate to repress the activation functions of E2F1. Proc. Natl. Acad. Sci. USA **94:** 11268–11273.
87. ZHANG, H.S. *et al.* 2000. Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF. Cell **101:** 79–89.
88. STROBER, B.E. *et al.* 1996. Functional interactions between the hBRM/hBRG1 transcriptional activators and the pRB family of proteins. Mol. Cell. Biol. **16:** 1576–1583.
89. DUNAIEF, J.L. *et al.* 1994. The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. Cell **79:** 119–130.
90. NG, H.H. *et al.* 1999. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. Nat. Genet. **23:** 58–61.
91. FUKS, F. *et al.* 2001. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. EMBO J. **20:** 2536–44.
92. DEPLUS, R. *et al.* 2002. Dnmt3L is a transcriptional repressor that recruits histone deacetylase. Nucleic Acids Res. **30:** 3831–3838.
93. MARKS, P.A., V.M. RICHON & R.A. RIFKIND. 2000. Histone deacetylase inhibitors: inducers of differentiation or apoptosis of transformed cells. J. Natl. Cancer Inst. **92:** 1210–1216.
94. JUNG, M. 2001. Inhibitors of histone deacetylase as new anticancer agents. Curr. Med. Chem. **8:** 1505–1511.
95. CHOPIN, V. *et al.* 2002. Sodium butyrate induces P53-independent, Fas-mediated apoptosis in MCF-7 human breast cancer cells. Br. J. Pharmacol. **135:** 79–86.
96. SAWA, H. *et al.* 2001. Histone deacetylase inhibitors such as sodium butyrate and trichostatin A induce apoptosis through an increase of the bcl-2-related protein Bad. Brain Tumor Pathol. **18:** 109–114.
97. HEROLD, C. *et al.* 2002. The histone-deacetylase inhibitor Trichostatin A blocks proliferation and triggers apoptotic programs in hepatoma cells. J. Hepatol. **36:** 233–240.
98. COHEN, L.A. *et al.* 1999. Chemoprevention of carcinogen-induced mammary tumorigenesis by the hybrid polar cytodifferentiation agent, suberanilohydroxamic acid (SAHA). Anticancer Res. **19:** 4999–5005.
99. COHEN, L.A. *et al.* 2002. Suberoylanilide hydroxamic acid (SAHA), a histone deacetylase inhibitor, suppresses the growth of carcinogen-induced mammary tumors. Anticancer Res. **22:** 1497–1504.
100. RICHON, V.M. *et al.* 1998. A class of hybrid polar inducers of transformed cell differentiation inhibits histone deacetylases. Proc. Natl. Acad. Sci. USA **95:** 3003–3007.

101. KIM, Y.B. *et al.* 1999. Oxamflatin is a novel antitumor compound that inhibits mammalian histone deacetylase. Oncogene **18:** 2461–2470.
102. YOSHIDA, M. & S. HORINOUCHI. 1999. Trichostatin and leptomycin. Inhibition of histone deacetylation and signal-dependent nuclear export. Ann. N.Y. Acad. Sci. **886:** 23–36.
103. KIM, M.S. *et al.* 2000. Apicidin, an inhibitor of histone deacetylase, prevents H-ras-induced invasive phenotype. Cancer Lett. **157:** 23–30.
104. HAN, J.W. *et al.* 2000. Apicidin, a histone deacetylase inhibitor, inhibits proliferation of tumor cells via induction of p21WAF1/Cip1 and gelsolin. Cancer Res. **60:** 6068–6074.
105. NAKAJIMA, H. *et al.* 1998. FR901228, a potent antitumor antibiotic, is a novel histone deacetylase inhibitor. Exp. Cell Res. **241:** 126–133.
106. JABOIN, J. *et al.* 2002. MS-27-275, an inhibitor of histone deacetylase, has marked in vitro and in vivo antitumor activity against pediatric solid tumors. Cancer Res. **62:** 6108–6115.
107. SAITO, A. *et al.* 1999. A synthetic inhibitor of histone deacetylase, MS-27-275, with marked in vivo antitumor activity against human tumors. Proc. Natl. Acad. Sci. USA **96:** 4592–4597.
108. MCCAFFREY, P.G. *et al.* 1997. Induction of gamma-globin by histone deacetylase inhibitors. Blood **90:** 2075–2083.
109. RANSOM, R.F. & J.D. WALTON. 1997. Histone hyperacetylation in maize in response to treatment with HC-toxin or infection by the filamentous fungus Cochliobolus carbonum. Plant Physiol. **115:** 1021–1027.
110. DEMARY, K., L. WONG & R.A. SPANJAARD. 2001. Effects of retinoic acid and sodium butyrate on gene expression, histone acetylation and inhibition of proliferation of melanoma cells. Cancer Lett. **163:** 103–107.
111. VIGUSHIN, D.M. *et al.* 2001. Trichostatin A is a histone deacetylase inhibitor with potent antitumor activity against breast cancer in vivo. Clin. Cancer Res. **7:** 971–976.
112. RASHID, S.F. *et al.* 2001. Synergistic growth inhibition of prostate cancer cells by 1 alpha,25 Dihydroxyvitamin D(3) and its 19-nor-hexafluoride analogs in combination with either sodium butyrate or trichostatin A. Oncogene **20:** 1860–1872.
113. HOU, M. *et al.* 2002. The histone deacetylase inhibitor trichostatin A derepresses the telomerase reverse transcriptase (hTERT) gene in human cells. Exp. Cell. Res. **274:** 25–34.
114. VAN LINT, C., S. EMILIANI & E. VERDIN. 1996. The expression of a small fraction of cellular genes is changed in response to histone hyperacetylation. Gene Expr. **5:** 245–253.
115. BUTLER, L.M. *et al.* 2002. The histone deacetylase inhibitor SAHA arrests cancer cell growth, up-regulates thioredoxin-binding protein-2, and down-regulates thioredoxin. Proc. Natl. Acad. Sci. USA **99:** 11700–11705.
116. MUNSTER, P.N. *et al.* 2001. The histone deacetylase inhibitor suberoylanilide hydroxamic acid induces differentiation of human breast cancer cells. Cancer Res. **61:** 8492–8497.
117. AMIN, H.M., S. SAEED & S. ALKAN. 2001. Histone deacetylase inhibitors induce caspase-dependent apoptosis and downregulation of daxx in acute promyelocytic leukaemia with t(15;17). Br. J. Haematol. **115:** 287–297.
118. DEROANNE, C.F. *et al.* 2002. Histone deacetylases inhibitors as anti-angiogenic agents altering vascular endothelial growth factor signaling. Oncogene **21:** 427–436.
119. YOSHIDA, M. *et al.* 2001. Histone deacetylase as a new target for cancer chemotherapy. Cancer Chemother. Pharmacol. **48** Suppl. 1: S20–26.
120. KWON, H.J., *et al.* 2002. Histone deacetylase inhibitor FK228 inhibits tumor angiogenesis. Int. J. Cancer. **97:** 290–296.
121. LEE, B.I. *et al.* 2001. MS-275, a histone deacetylase inhibitor, selectively induces transforming growth factor beta type II receptor expression in human breast cancer cells. Cancer Res. **61:** 931–934.
122. COTE, S. *et al.* 2002. Response to histone deacetylase inhibition of novel PML/RARalpha mutants detected in retinoic acid-resistant APL cells. Blood. **100:** 2586–2596.
123. FERRARA, F.F. *et al.* 2001. Histone deacetylase-targeted treatment restores retinoic acid signaling and differentiation in acute myeloid leukemia. Cancer Res. **61:** 2–7.

124. HE, L.Z. *et al.* 2001. Histone deacetylase inhibitors induce remission in transgenic models of therapy-resistant acute promyelocytic leukemia. J. Clin. Invest. **108:** 1321–1330.
125. KITAMURA, K. *et al.* 2000. Histone deacetylase inhibitor but not arsenic trioxide differentiates acute promyelocytic leukaemia cells with t(11;17) in combination with all-trans retinoic acid. Br. J. Haematol. **108:** 696–702.
126. LIN, R.J. *et al.* 1998. Role of the histone deacetylase complex in acute promyelocytic leukaemia. Nature **391:** 811–814.
127. MINUCCI, S. *et al.* 2001. Histone deacetylases: a common molecular target for differentiation treatment of acute myeloid leukemias? Oncogene **20:** 3110–3115.
128. PANDOLFI, P.P. 2001. Transcription therapy for cancer. Oncogene **20:** 3116–3127.
129. FINNIN, M.S. *et al.* 1999. Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. Nature **401:** 188–193.
130. GROZINGER, C.M. & S.L. SCHREIBER. 2002. Deacetylase enzymes: biological functions and the use of small-molecule inhibitors. Chem. Biol. **9:** 3–16.
131. RANDHAWA, G.S. *et al.* 1998. Identification and mapping of human histone acetylation modifier gene homologues. Genomics **51:** 262–269.
132. FURUKAWA, Y. *et al.* 1996. Isolation and mapping of a human gene (RPD3L1) that is homologous to RPD3, a transcription factor in Saccharomyces cerevisiae. Cytogenet. Cell Genet. **73:** 130–133.
133. BETZ, R. *et al.* 1998. Human histone deacetylase 2, HDAC2 (Human RPD3), is localized to 6q21 by radiation hybrid mapping. Genomics **52:** 245–246.
134. WANG, A.H. *et al.* 1999. HDAC4, a human histone deacetylase related to yeast HDA1, is a transcriptional corepressor. Mol. Cell. Biol. **19:** 7816–7827.
135. TONG, J.J. *et al.* 2002. Identification of HDAC10, a novel class II human histone deacetylase containing a leucine-rich domain. Nucleic Acids Res. **30:** 1114–1123.
136. GUARDIOLA, A.R. & T.P. YAO. 2002. Molecular cloning and characterization of a novel histone deacetylase HDAC10. J. Biol. Chem. **277:** 3350–3356.
137. KAO, H.Y. *et al.* 2002. Isolation and characterization of mammalian HDAC10, a novel histone deacetylase. J. Biol. Chem. **277:** 187–193.
138. FRYE, R.A. 2000. Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. Biochem. Biophys. Res. Commun. **273:** 793–798.
139. NEWMARK, H.L. & C.W. YOUNG. 1995. Butyrate and phenylacetate as differentiating agents: practical problems and opportunities. J. Cell. Biochem. Suppl. **22:** 247–253.